Vowel harmony in isiXhosa: an OT and acoustic study of [ATR]

A thesis submitted in fulfilment of the requirements for the degree of

MASTERS IN ENGLISH LANGUAGE & LINGUISTICS

of

RHODES UNIVERSITY

By KELLY KILIAN

December 2017

Abstract

The vowel harmony system in isiXhosa is centred on a process of vowel raising. All mid-vowels preceding a high vowel take on the feature advanced tongue root (ATR) (e.g. thɛnga ‘buy’ → thengisa ‘sell’; bɔna 'see' → bonisa 'cause to see') (Harris 1987). The process of mid-vowel assimilation for the feature [+ATR] is consistent in all instances in which the mid-vowel occurs preceding a high vowel trigger, unless harmony is blocked by the low opaque vowel [a]. This is the analysis presented in Jokweni & Thipa (1996) the only previous literature to address the vowel harmony process of isiXhosa in detail. As an alternative approach to the rule-based phonology applied in the analysis presented by Jokweni & Thipa (1996), I propose the introduction of Optimality theory (OT) (Prince & Smolensky 1993, Bakovic 2000, and Pulleyblank 2002). I will present a map of the harmony system of isiXhosa using OT, while also presenting acoustic data to supplement the selected examples provided in Jokweni & Thipa (1996). This acoustic investigation will determine whether the harmonic feature is ATR, and how this feature patterns among vowels in different phonological contexts.

In this paper vowel harmony is achieved through the implication of numerous rules, and with very specific directional and prosodic limitations on the spread of [+ATR]. Using generalisations based on my own collected data as well as those reported in previous literature, I have developed a constraint ranking to account for the harmony process in isiXhosa. By adapting the No-disagreement approach to harmony (Pulleyblank 2002), the final constraint ranking has the capacity to derive the optimal phonetic candidate for every harmony case. A selection of spread constraints is used to account for the raising as well as blocking processes, by driving either regressive or progressive spreading. Within the original No-disagreement approach a spread constraint would recognised only one feature in its prohibition of disagreeing segments. However, in the adapted approach the spread constraint driving [+ATR] assimilation is combined with a feature of correspondence (Krämer 2001) which considers the height as well as the ATR value of the sequential segments. The constraint is therefore adapted to consider more than one feature and is not activated unless the sequential segments agree for this particular feature. The regressive spread constraint is therefore only activated when the consecutive segments have an agreeing height value. The introduction of this adaptation was necessary to provide a more nuanced OT approach with the capacity to effectively characterise the idiosyncrasies observed in this harmony pattern.

The harmony constraints are therefore no longer contradict one another by simultaneously driving harmony in opposite directions.

Furthermore, the direct acoustic analysis is completed by means of the PRAAT software, to answer the salient question of the definitive harmonic feature. To provide a multiplicity of empirical evidence I have recorded utterances containing a number of vowel combinations. Each combination positions the alternating mid-vowels in a particular phonological context from which instances of ATR alternations have been extracted and phonetically analysed. Using the generalisations reported in Jokweni & Thipa (1996) as a starting point, the acoustic signal of each mid-vowel within a set phonological context is annotated for a predicted ATR value. Hence, if a mid-vowel occurs preceding a high vowel it is annotated as [+ATR] etc. The data sets representing each of the mid-variants found in a specific phonological context are then plotted into vowel charts and compared by means of statistical analysis (Baayen 2008, Bluman 2000). The results are then used to determine whether any significant phonetic alternation is occurring, and what the acoustic distinction between [+ATR] & [-ATR] variants is essentially comprised of. The final acoustic results indicate a significant difference between the mid-vowel ATR variants extracted from specific phonological contexts. Hence, due to co-articulatory effects or some other phonological influence the realisation of [+/-ATR] variants exist along a spectrum, and are therefore not phonetically consistent, but indicate a different acoustic make-up across the various groups.

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor Dr William Bennett of the English Language & Linguistics Department of Rhodes University. Most of our correspondence was conducted via Skype, as Dr Bennett moved to the University of Calgary while I was working on my thesis. However, whether in person during his visits or via Skype, I always had the support I needed to produce this research. Throughout every step of the process he consistently pushed me to deliver work that is well researched and probing. Without his supervision and encouragement this investigation would not have been successful. Secondly, I would like to thank my informant Olwethu Zeleni, who spent hours producing the utterances which make up my collected data. Of course, I would like to thank the lecturers of the English language & Linguistics Department of Rhodes University. All of whom have been essential to my appreciation and understanding of linguistic studies. Without their assistance and guidance throughout my undergraduate studies I may not have felt equip to pursue post graduate studies in linguistics. Finally, I would like to thank my amazing parents James & Lynette Kilian, for providing me the opportunity of a tertiary education at a prestigious institution like Rhodes University. Without their encouragement and sacrifice this level of accomplishment would not have been possible. In my experience, the most challenging part of producing a thesis has been "The notion of sitting down and conjuring up, not only words in which to clothe thoughts but thoughts worthy of being clothed" (Fitzgerald 2009)

Table of Contents

Table of Figures

1. Introduction

1.1. Vowel harmony in isiXhosa

A system of vowel harmony has been reported on in isiXhosa. The type of harmony is analysed as advanced tongue root (ATR) harmony (Jokweni & Thipa 1996). In the previous analysis conducted by Jokweni & Thipa (1996), featural and structural observations of the pattern are presented in the form of generalisations as well as a final structural rule. The vowel inventory of isiXhosa consists of 7 surface vowels, 5 of which exist underlyingly and are considered inherent to the system (Pahl 1967; see also Harris 1987, Jordan 1966, Sibanda 2009).

(1) Underlying vowels:

	front		back
+ATR	**i**		**u**
-ATR	**ɛ**		**ɔ**
-ATR		**a**	

Non-phonemic vowels:

	front	back
+ATR	**e**	**o**

As presented in the tableau above the mid-vowels are the only segments which have harmonic counter parts for ATR. The advanced set [e & o] resemble the orthographic mid-vowel representations and are also considered the marked variant (Harris 1987). Hence, the basic pattern states that the spread of [+ATR] ,and therefore the non-phonemic set of mid-variants will only occur when vowel raising is triggered by a high vowel [i/u] (Jokweni & Thipa 1996). As the [+ATR] variants have not been proven to exist underlying there is a non-contrastive relationship between these mid-vowel sets distinguished for ATR. Hence, in a phonological context in which a high vowel triggers the raising of a preceding mid-vowel, the mid-variant will undergo a phonetic alternation, however the semantic content of the word will remain the same. The collected generalisations needed to define this harmonic system may be categorised into two sets. These include the function of individual vowel segments within the system such as the trigger function of the high vowel, and secondly the structural idiosyncrasies of the pattern itself. Beginning with the examples below, the transcription convention followed throughout this paper include all vowels presented in IPA format, and all consonants in orthographic form.

(2) Vowel behaviors:

a. Mid-vowels undergo harmony
b. High vowels trigger harmony
c. Low vowel blocks raising

(3) Structural generalisations:

a. [-ATR] is the default feature
b. Regressive spreading
c. The rule of [+ATR] spreading is iterative

(4) Illustrative examples

a. Mid-vowels undergo harmony:

/-ɛnza/ 'make, do'

i)	ndɛnza	1st.sbj, pres
ii)	wɛnza	2nd.sbj, pres
iii)	ndɛnzɛ	1st.sbj, perf
iv)	wenziswa	2nd.sbj, caus, pres
v)	ndenzilɛ	1st.sbj, perf.disj

b. High vowels trigger harmony

/-qɔnda/ 'understand'

i)	ndiyaqɔnda	1st.sbj, pres
ii)	uyaqɔnda	2nd.sbj, pres
iii)	uyaqɔndɛ	1st.sbj, perf
iv)	uyaqondiswa	2nd.sbj, caus, pres
v)	uyaqondilɛ	2nd.sbj, perf.disj

c. Low vowel blocks raising

i)	unɛsandlozi	'his/her wide eyes starring'
ii)	ɛbhalekiswa	's/he is caused to run'
iii)	ɛbhakontini	'outdoor brick oven'

To summarize, mid-vowels are the only vowels which undergo harmony as displayed in the examples in set (a). The mid-vowels are therefore the only vowels which alternate in different phonological contexts. In set (b) high vowels triggering raising is demonstrated, as the advanced mid-vowel form only occurs in the causative and initial past tense constructions. Finally, low vowel blocking is demonstrated in set (c) as all mid-variants preceding the low vowel occur in the retracted form. The generalisations above describe the function of the different vowels in the harmony system, the generalisations to follow describe the structural specifications of the system itself.

(5) Structural generalisations:

a. [-ATR] is the default feature

i) uqwɛthɛ qwɛthɛ 'irritation in the mouth or throat'

ii) uqhɔqhɔqhɔ 'windpipe/trachea'

iii) umɛndɔ 'high road'

b. Regressive spreading

i) nditengilɛ 'buy'

ii) ndifeketilɛ 'play/trifle'

iii) ndisebenzilɛ 'work'

c. The rule of [+ATR] spreading is iterative

i) enthlekeleni 'in a calamity'

ii) esikolweni 'at school'

iii) esekenziswa 'caused to work'

To summarize, in example set (a) [-ATR] is seen to be the default form as it is the variant which occurs when there are no high or low vowel influences. A particular directionality is exhibited in set (b) as all mid-variant preceding a high vowel are triggered for raising, while those following occur in the retracted form. The final example set demonstrates the iterative nature of the spread rule as it continues across morphological and phonological boundaries in these instances.

1.2. Analysis

The aim of this research is to analyse the harmony system in OT, and furthermore to provide substantial empirical data of the phenomenon. Only having been reported on once by Jokweni

& Thipa (1996), the phenomenon exhibited above still requires extensive investigation. Some of the generalisations above are only slightly adapted from those presented in Jokweni & Thipa (1996). All evidence collected in support of the phenomenon consisted of a very limited set of examples of each occurrence, which were manually transcribed. Hence, the basis of the generalisations in the previous literature is comprised of individual transcription by researchers, based solely on their own perceptions of the mid-vowel variants. These observations were stated in the form of generalisations and adapted to fit a parametrical rule which accounts for the structural necessities required to derive harmony. Jokweni & Thipa (1996) attempted to apply a contemporary theoretical approach to this African language phenomenon, which they achieve in the form of this parametrical rule. The questions left open however, require an acoustic analysis of the variants reported as [+-ATR] as in the investigation of mid-vowel assimilation in SiSwati (Malambe 2015). This form of analysis will provide empirical evidence which will determine whether this phenomenon does exist, and whether it is ATR. In an attempt to continue applying novel and contemporary theories to research in African languages the salient part of the analysis conducted in this paper uses Optimality theory to account for the vowel harmony process. The generalisations presented in chapter three are hence revisited with original data collected from a mother tongue speaker.

1.3. Optimality Theory

Many contemporary analyses of the vowel harmony phenomenon have applied Optimality Theory (OT) as a form of analysis (Archangeli & Pulleyblank 1989, Bakovic 2000, Beckman 1997, Pulleyblank 2002, Sasa 2009). A uniform approach capable of dealing with the idiosyncrasies that exist cross linguistically is however yet to be developed. There are numerous approaches to OT, each designed to derive harmony with focus on one dominant aspect of the system. Approaches have been favoured across previous investigations of vowel harmony, these include span theory, ABC harmony, and spread theory. Each with their own postulations and predetermined axioms, these approaches have successfully accounted for different harmony cases. There are however languages with more complex harmony systems which cannot be accurately characterised and accounted for by these approaches (Sasa 2009). The capability of these approaches is limited first as result of their prior postulations but furthermore because they lack the capacity to account for both issues of directionality as well as sour grapes (Wilson 2006).

Having considered and discussed the possible application of OT to the vowel harmony process in isiXhosa, an adaptation of the no-disagreement approach (Pulleyblank 2002) is presented as the most effective option. The no-disagreement approach unlike those mentioned above requires no prior postulations to function effectively, and hence could be described as the most transparent and uniform option. An adaptation was necessary however to account for the default requirement, the regressive directionality and finally the issue of sour grapes. As isiXhosa has an extremely complex harmony system the OT approach applied must have the capacity to account for each of these issues simultaneously. In order to successfully characterise this particular harmony process, an aspect of correspondence theory was adopted and combined with the agreement constraint. The addition of this element enables the agreement constraint to consider more than one feature, and therefore will only activate when faced with the proximity of segments already retaining a particular feature. Without this constraint a final constraint ranking cannot be produced in the OT workplace program. The basic no-disagreement constraint ranking results in several contradictory ercs. The main contradiction arises in the relationship between the harmony constraints which produce blocking and raising. Each harmony constraint either requires too much regressive spreading or too much progressive spreading, the ranking is not capable of accounting for the nuances of this process. The adaptation of this approach is however successful in deriving the attested form in every harmony case.

The analysis is structured to begin with the adapted form of the no-disagreement theory applied to specific harmony cases. isiXhosa has instances of total harmony, forms of partial harmony as well as instances in which no harmony is derived. The adapted theory is tested against each of these cases to determine whether the approach is capable of consistently deriving the attested form. This is then followed by a demonstration of the basic no-disagreement theory applied to isiXhosa. This includes particular focus on the specific constructions and ercs revealed to be problematic by OT workplace. The following discussion then focusses on further elucidating the issues concerning the theory and the specificities of this harmony process. Thereby simultaneously providing justification for the adaptation to the theory.

1.4. Acoustic analysis

The harmonic feature Advanced tongue root is generally described as the extension and reduction of the pharyngeal cavity (Lindau 1978). Any change in this type of dorsal constriction is accomplished mainly by a shift in the position of the tongue and jaw. When the tongue root is advanced, the pharyngeal cavity is extended, when retracted the cavity is reduced. These changes in the positioning of the tongue, and shape of the oral cavity are reflected in the formant frequencies. The first and second formants are considered most sensitive to these articulatory changes, and their measurement values therefore generally reflect any type of alternation (Reetz & Jongman 2009). Using PRAAT software the F1 & F2 values of mid-variants in particular phonetic contexts were measured and extracted. As explained in the generalisations above, depending on the phonetic context the mid-variant would either occur as the advanced or retracted form. Using novel recorded data each vowel is annotated as its predicted ATR value based on the phonological context. Comparisons between formant resonant values of [+ATR] & [-ATR] mid-variants indicate that there is a significant acoustic difference between these harmonic sets. Hence, there is some type of harmony process producing these alternations. The investigation of the F1 & F2 resonant values between different variants then further systematically point to ATR as the definite harmonic feature.

The structure of this analysis begins with the presentation of the vowel charts. The plots illustrate the position of each advanced and retracted variant in relation to each other, as well as other segments within the vowel space. Each chart represents a specific comparison or mid-variant, which is immediately followed by a description of the plot and the salient variant on display. There is a total of 6 vowel charts presented in this chapter, each of which function to answer a specific question concerning the acoustic significance of this phenomenon. An analysis based solely on these mappings would consider all discrepancies between advanced and retracted variants to be significant. Hence, a statistical analysis is also conducted, specifically a two-way ANOVA with replication. Following each vowel chart are one or two ANOVA tableaux. These statistical calculations are presented directly following the vowel charts to allow for comparison and the interpretation of significant discrepancies.

2. Theoretical background

2.1. Vowel Harmony

Vowel harmony is the systematic interaction between vowels within a specific domain. As most languages use a CVC (consonant, vowel, consonant) structure vowel harmony is used to describe the relation between nonadjacent segments, as the syntactic structure limits the adjacent occurrence of vowels. An understanding of auto-segmental phonology is integral to the analysis of vowel harmony, as it enables the analysis of sounds as they relate on different tiers and in different places. However, contemporary studies of harmony patterns rely heavily on some form of acoustic analysis as well as various approaches to optimality theory. Through the study and analysis of vowel harmony in different languages we are able to identify the features which dictate agreement, as well as those which tolerate disharmony between vowel segments within the relevant domains. The description of the vowel harmony system of a language can be used to further understand the typology of that language, which will enable the further categorisation and classification of world languages (Van de Weijer 1996).

2.1.1. ATR harmony

Advanced tongue root harmony between vowel segments, is generally understood as an agreement between segments in relation to the size of the pharynx during articulation. This type of harmony is commonly found in Niger-Congo languages, Nilo-Saharan families, as well as numerous Bantu languages (Williamson 1984). Languages of this harmony type drive segments to share the same value for ATR. Hence, aside from exceptional circumstances there should be no constructions with both [-ATR] and [+ATR] segments occurring in a particular domain (Casali 2008). Although found in languages from other geographic regions, this type of harmony and this particular harmonic feature is common across the harmony systems of African languages. In the most straightforward case the vowel inventory will consist of 10 vowels with equal harmonic counter parts for ATR, in languages such as Bongo; Degema; Diola-Fogny; Koromfe and Vata. In these symmetric cases it is much less likely to find the occurrence of opposing ATR vowel segments within the same root, affix, suffix or prosodic word. The most common alternations occur in the form of a harmonizing affix, as opposed to an alternating root. Hence, if the word-root contains a [-ATR] segment any vowel segments in the attached affix/s should be changed to the [-ATR] form. Examples of this type of alternation can be found in languages such as Kasem spoken in Ghana and Burkina Faso (Casali 2008). A symmetrical vowel inventory of this nature is however quite uncommon, and many systems of 9, 8, or even 7 vowel segments are found. Without harmonic counterparts for each vowel, it becomes challenging to consistently alternate to a particular ATR counterpart. Hence, there are numerous cases in which irregular harmony patterns are found. These are considered general limitations of ATR harmony and may manifest as the failure of particular affixes to undergo harmony due to their idiosyncratic nature, or to the scope of harmony limited to a specific morphological or phonological domain (Archangeli & Pulleyblank 2007).

In many of the asymmetrical vowel inventories the [-ATR] low vowel [a] lacks a harmonic counterpart and is permitted to co-occur with [+ATR] vowel segments. Although certain languages may contain the same number of vowels in their inventory for example there are numerous 7 vowel systems, the vowel segments which exist within the inventory may be different. Hence, in Kinande the vowel inventory lacks the [+ATR] mid-vowel forms e & o (Mutaka 1995), while the inventory of Yoruba lacks the [-ATR] high vowel forms ɪ & ʊ (Archangeli & Pulleyblank 1989). Particular vowel inventories found in the West Atlantic language Pulaar as well as the Bantu languages Tsonga, isiZulu as well as isiXhosa, underlying contain only 5 vowels, however a phonetic contrast may be found on the surface with two sets of mid-vowels distinguished for ATR. Such systems are considered to have a phonemic 5 vowel system, while simultaneously having a phonetic 7 vowel system. In these harmony systems ATR is not contrastive but rather represents an allophonic alternation, involving the assimilation of the mid-vowel to a particular [+ATR] or [-ATR] vowel segment. The more robust harmony cases involving ATR are of course those in which the harmonic feature is recognised underlying on a phonemic level and serves a contrastive function semantically (Casali 2008).

2.2. OT an introduction

The constraints and constraint ranking are considered the drivers in the analysis of the vowel harmony process. These constraints are used to limit and elicit the correct surface forms which occur in a particular language. The requirements stipulated by each constraint ensure that the optimal candidate in each case is the candidate that appears in the surface form. The choice of theory is therefore based on the preconditions of the language as well as the relationship between the underlying and the surface forms. There are a number of possible approaches and methods in which constraints are utilised, these involve different functions and more importantly place focus on different aspects of the harmony process. The basic definition of the harmony process, as described in derivational autosegmental theory (Clements 1985) is that lexical representations within a language will choose between specifications for different features. Secondly, this choice or specification will be required throughout a specific domain as a result of spreading. In its most basic occurrence, vowel harmony appears to be the assimilation of one feature or feature value, across a selection of anchors (Pulleyblank 2002). As complete harmony does not occur in every language, the process of assimilation must be tailored to accommodate the harmony pattern of a particular language.

2.2.1. Constraint types

The advent of Optimality Theory in the analysis of vowel harmony has resulted in an array of differing approaches, making the challenge of analysis even more complex as the choice of constraint and ranking must be justified in the resulting harmony pattern. As mentioned above, the functions of differing approaches to harmony within OT shift the focus of the analysis to different aspects of the harmony process. Hence, alignment constraints focus on placing limitations on both the domain, as well as the directionality of any assimilation. This requires all anchors from the left to right edges of a domain be aligned for a specific feature. This constraint might take the form: Align(F)L, which performs the function align for the feature F to the left edge of the prosodic word or domain (Kirchner 1993). If a language has either a distinct progressive or regressive spreading pattern, this would be easily accounted for by an alignment constraint like the one above. It does however serve as problematic if the harmonic domain is designated within a morphological or phonological context within the prosodic word. The optimal domains theory (Cole & Kisseberth 1994) then elevates the status of the domain, requiring the construction of abstract harmonic domains as a result of alignment. Hence, if a domain cannot be defined either phonologically or morphologically, this approach will not function at its full capacity.

Possibly the most productive constraint is the agreement constraint. This requires that all adjacent segments share the same value for a specific feature. However, other features influencing the harmony pattern may also be encoded in such a constraint: Spread-L([F], D) this constraint can be described as agree to the left for the feature F in domain D. Featural prohibitions and positional faithfulness place limitations on which feature value is permitted on the surface, as well as the relationship between the input form and the lexical representation. If a specific value for a feature does not appear in the surface form within a language then it would be useful to prohibit the occurrence of this feature value completely, or simply limit its occurrence to certain positions within the prosodic word by restricting the positions in which segments may alternate. One might therefore use the following constraint (*+F >> *−F >> IDENT[F]) to prohibit the feature value +F, ranking this above a prohibition on the feature value -F will result in the feature –F appearing in the surface form (Beckman 1997). As seen in the constraint outline above, these prohibitions are ranked above any faithfulness constraints to ensure that the harmony process may occur. Hence, there are a variety of OT approaches which are constructed around a set of constraints that function in a particular way. As discussed above, particular constraints are more effective than others in analysing different harmony patterns. The ultimate goal is therefore to identify an approach to OT which can be effectively applied to every type of harmony case in a uniform system.

2.2.2. Types of Harmony patterns

The selection of constraints and overall approach used to analyse the phenomenon of vowel harmony in a particular language must be explained by the surface forms of that language as well as the suboptimal candidates. The appropriate OT approach must be capable of accounting for the diverse harmony patterns that occur cross linguistically. To discuss some of the situations in which certain OT approaches are lacking either in empirical or theoretical grist the possible harmonic processes must be described. Cross-linguistically the harmony patterns found may be placed into three general categories. It may be the case that all vowels within a domain such as the prosodic word all share the same feature, in which case this is referred to as total harmony. There are also cases in which harmony occurs across a medial vowel which holds a different value for the same feature, this is referred to as transparency. In such cases the medial vowel is considered transparent and therefore does not participate in the harmonic process. The final category contains the pattern in which only one vowel does not share the same value for a specific feature, however the two adjacent vowels agree for a specific feature value, this is referred to as opacity. Each pattern is presented schematically below (Sasa 2009):

(1) Vowel harmony types, schematically

Input /αF…βF…αF…βF/	V1	V2	V3	V4
Total harmony	[αF]	[αF]	[αF]	[αF]
Transparency	[αF]	[βF]	[αF]	[αF]

Opacity	[αF]	[αF]	[βF]	[βF]

The complications in opting for a particular OT approach arise in providing a comprehensive account of each of the patterns presented above. The analysis must account for cases of total harmony, partial harmony as well as no harmony. Referring to an investigation into the capabilities of three of the OT approaches (Sasa 2009), the implications of their selection will be discussed below. As explicated above, the alignment constraint evaluates candidates according to whether the anchors at the left and right edges of the word share the same feature value. Hence, the alignment constraint is less concerned with the medial vowel and therefore cannot accurately categorise the partial harmony process. This leaves one with a constraint ranking which has difficulty dealing with unattested patterns and is unable to theoretically characterise the partial harmony pattern comprehensively. Another important complexity of this constraint family is the approach to evaluation. The alignment approach evaluates candidates gradiently, which counts the number of vowels by which the harmonic feature misses the edge of the domain. Again this focus on harmony at the edge does not prefer the harmonicity of the highest number of vowels, but simply the proximity of the harmonic feature to the edge. While the alignment constraint favours transparency, the spread or feature linking and agreement constraints prefer opacity. These constraint types evaluate the violations according to adjacent assimilation within a specific domain, as oppose to agreement at the edge.

2.3. Optimality Theory applied to harmony in isiXhosa

2.3.1. Harmony by no-disagreement

The first OT approach to be applied to the harmony case of isiXhosa is Pulleyblank's harmony drivers (2002). This approach focuses on achieving harmony by prohibiting disharmony. Therefore, the general rule is that segments which hold opposite values for the same feature should be prohibited from occurring sequentially (Smolensky 1993). As discussed above there has been little consensus on the optimal method for deriving harmony within the OT system of analysis. This approach to OT focuses on driving harmony using a sample set of constraints targeting sequential prohibitions in order to achieve agreement. The usefulness of each constraint must be exemplified in the exigencies of the pattern. Hence, the functional motivation of each constraint must be consistently questioned and tested. Furthermore, the formulation of any new constraint must reflect its function in producing the appropriate harmony pattern. In adopting the no-disagreement approach, the constraints driving harmony are unified with the obligatory contour principle. The harmony constraints within this approach state that two segments may not disagree for a particular feature, the OCP as adapted and presented in Suzuki (1998) places a prohibition on sequences of elements in general. By aligning this OT approach with the OCP in terms of the process through which harmony is produced, the OT approach becomes more naturalised and intuitive.

2.3.2. Harmony & the OCP

Since the advent of OT it has become more and more challenging to identify a uniform approach to the analysis of vowel harmony. There are various approaches within the parameters of optimality theory including the no-disagreement approach, which aligns itself with the Obligatory Contour Principle (OCP) by enforcing the same basic featural requirements on sequential segments. A generalised description of the OCP focuses on the prohibition of specific sequences of segments (Suzuki 1998). Similar to the harmony constraint which forms the basis of the no-disagreement ranking, the OCP requires that a sequence of features be prohibited e.g. *X…Y, the sequence X, Y cannot occur in the attested form. The interaction between these segments is also dependent on their proximity and similarity between the elements. Hence, the interaction becomes stronger when the elements in question are adjacent or closer vs. when they occur at a distance. As well as when the segments share opposite values for the same element, this forces a stronger interaction. According to an understanding of this relation between the OCP and the no-disagreement approach, a uniform analysis of the vowel harmony system is made possible. It is proposed that even the effects of proximity are visible in the application of harmony constraints. Generally one finds the more strict or local assimilation is produced by the local harmony constraint e.g. *XY an adjacent xy sequence is prohibited. While patterns with cases of opacity prefer the sequential prohibition to be at a medium distance rather than applying to strictly adjacent segments. Finally the harmony patterns reflecting transparency in certain cases prefer the distance harmony constraint as it allows for disharmony between segments.

2.3.3. Opacity vs. transparency: proximal vs. distant featural prohibitions

2.3.3.1. Case studies: Yoruba & C'Lela

2.3.3.1.1. Introduction: Background

The No-disagreement approach is applied to the harmony case study of the language Degema (Pulleyblank 2002) which has a tongue root harmony system with a root controlled assimilatory pattern. Therefore, the root may either appear as advanced or retracted, while the affixes assimilate. In order to achieve this specific harmony pattern, the harmony constraints *ATR-co-RTR, *RTR-co-ATR must be ranked above the faithfulness constraints in order to drive harmony. Within the faithfulness constraints a domain or positional faithfulness must be specified. As the pattern exhibits a case of root-controlled harmony, the faithfulness constraint MAX(ATR)/root must be ranked above MAX(ATR)/word. Hence, while the harmony constraints work to prohibit segments of opposite features the positional faithfulness constraints retain the harmonic feature of the non-alternating segments, the segments which serve as the anchor. Based on this ranking, neither of the harmonic features is considered the marked option or the default. Therefore, it is imperative that the harmony constraints be ranked above all faithfulness constraints, however the order in which these constraints are ranked in relation to each other is not important and does not affect the outcome. Having positional faithfulness to the root outranking general faithfulness to the input, reflects the anchored status of the vowels in the root, and the alternating nature of the affix vowels. This harmony pattern is therefore successfully and accurately characterised by this approach.

(2) Harmony in Degema:

/U-[der]-AM/	*ATR-C-RTR	*RTR-C-ATR	MAX(ATR)/root	MAX(ATR)/word
a. [uderəm] ☺				
b. [uderam]	*			
c. [ʊderəm]		*		
d. [ʊderam]	*	*		
e. [ʊdɛram]			*	*

The tableau above extracted from Pulleyblank's Harmony drivers (2002), selects an input form in which only the vowel within the root has an underlying ATR specification. The ATR value of the affix vowels are therefore completely reliant on the constraint ranking. The initial losing candidates (2b) & (2d) violate the first no-disagreement constraint, which prohibits the occurrence of the feature ATR followed by RTR. Both of these candidates have an RTR low vowel affix following the ATR root vowel. The next constraint which prohibits the opposite featural sequence, no RTR segments followed by ATR segments is violated by candidates (2c) & (2d). This eliminates candidate (2c) which has a prefix RTR vowel, and therefore does not have the necessary agreement between the root and suffix as defined in the constraint. The final candidates (2a) & (2e) are both examples of total harmony, candidate (2a) reflects total ATR harmony, while candidate (2e) has total RTR harmony. The deciding constraint here assigns a violation on candidate (2e) due to the featural change of the root vowel, which is underlying specified as ATR. According to this ranking featural disagreement is not tolerated with either sequence prioritised. Therefore, although total harmony is favoured, the ranking requires that the shared feature be in agreement with the underlying feature specification of the vowel within the word root.

The capacity of this approach to deal with patterns of opacity and transparency are clearly exhibited in the case studies of Oyo & Ife, two dialects of Yoruba. The standard variety of Yoruba generally has a seven vowel system {i, e, ɛ, a, ɔ, o, u}. The harmonic feature identified in the harmony pattern is the position of the tongue root ATR. In both standard Yoruba as well as Oyo the high vowels act as opaque, while in Ife the high vowels are transparent. Hence, we have a case of mid-vowel tongue root assimilation as in the Degema pattern, however disharmony is forced in the attempt to retain the tongue root specification of the root while adhering to requirements of retraction and advancement. In lexical representations with only mid-vowel roots, full harmony is achieved with the same basic constraint ranking as Degema. The cases of interest however are those lexical representations containing high vowels. Specifically the cases in which a high vowel is situated between two mid-vowels. The case in which the high vowel is opaque forces the preceding mid-vowel to agree for the feature ATR, while only the final mid-vowel can occur in the retracted form. In cases of transparency however the mid-vowels on either end will harmonise with each other as oppose to assimilation to the high vowel. In order to derive such patterns of opacity as attested in the varieties of standard Yoruba and Oyo and then transparency in Ife, the distance of sequential prohibition must be edited.

2.3.3.1.2. Opacity pattern of Oyo (Pulleyblank 2002):

(3)

/ElUbO/	*HI/RTR	MAX(RTR)/root	*RTR-C-ATR	*ATR-C-RTR
a. [elubɔ] ☺				*
b. [ɛlʊbɔ]	*			
c. [ɛlubɔ]			*	*
d. [elubo]		*		
e. [ɛlubo]		*	*	

The tableau above illustrates the no-disagreement constraint ranking needed to produce the opaque harmony pattern in both standard Yoruba & the dialect Oyo. As in Degema, the root vowel in the Yoruba dialects are specified underlying for either ATR or RTR, therefore the prefix in the input form contains two vowels which will be assigned a tongue-root specification as determined by the constraint ranking. Candidate (3b) is the first to be eliminated as it violates the highest ranked constraint *HI/RTR by including a high vowel with the feature RTR, which is not included in the vowel inventory of Yoruba. Candidates (3d) & (3e) are eliminated because the root vowel specified as RTR in the input form appears as ATR. The first harmony constraint which considers the features of vowel segments only separated by one consonant, is violated by candidates (3c) & (3e) which both have prefixes with an RTR segment followed by a high vowel retaining the feature ATR. The optimal candidate is candidate (3a), which is only assigned a violation by the final harmony constraint prohibiting an ATR segment to be followed an RTR segment. The attested candidate is therefore one which prohibits the occurrence of an RTR segment, except for the final vowel.

While this ranking is able to derive opacity, some adjustments must be made to account for cases of transparency as in Ife. As proposed initially, a distance form of the sequential prohibition constraint is needed in order to produce transparency. The original harmony constraints *RTR-C-ATR & *ATR-C-RTR are therefore replaced with the distance forms *RTR~ATR & ATR~RTR.

2.3.3.1.3. Transparency pattern of Ife (Pulleyblank 2002):

(4)

/ElUbO/	*HI/RTR	MAX(RTR)/root	*RTR-∞-RTR	*ATR-∞-ATR
a. [elubɔ]			*	**

b. [ɛlʊbɔ]	*			
c. [ɛlubɔ] ☺			*	*
d. [elubo]		*		
e. [ɛlubo]		*		

The OT analysis above produces the transparent harmony pattern attested in Ife. Candidates (4b), (4d) & (4e) are eliminated in the same way as those in Oyo by the first two constraints, which prevent any high RTR vowel segments or any root vowels alternating from to ATR. Candidate (14c) the optimal candidate in Oyo is eliminated by the distance harmony constraint prohibiting ATR segments to be followed by RTR segments. This constraint assigns violations to both candidates (4a) & (4c), however it is the consideration of every segment within the now long-distance environment that distinguishes between the optimal and sub-optimal candidate. The long-distance constraint is therefore what determines the preference for transparency, as it assigns more violations to candidate (14a) which is the attested form in Oyo.

2.4. Review of alternative approaches to vowel harmony in OT

The recently developed approaches to OT include a modernised account of the feature spread constraint, as well as the development of span theory and the agreement by correspondence analysis. The introduction of these approaches became necessary when none of the existing approaches could deal with certain empirical issues, in particular cases of partial harmony. This is evidenced in the case of the Niger-Congo language Pulaar, for which none of the approaches account for the issues of directionality and sour grapes (Padgett 1997, Wilson 2006). Directionality here refers to the pattern of regressive spreading in Pulaar vowel harmony (Archangeli & Pulleyblank 1994, Sasa 2009). The sour grapes concept refers to the situation in which candidates with no harmony are selected as optimal above those with partial harmony, in the cases in which a blocker is present.

(5) Sour grapes:

/barɔdi/	*ɪ/ʊ	Spread [ATR]	Id (Root) [ATR]	Id [ATR]
a) barodi		***	*!	*
b) barɔdi ☺		***		
c) barɔdɪ	*!			*

This concept therefore refers to the problem in which the process of harmony assimilation is predicted either as total harmony or no harmony, eliminating the possibility of partial harmony (Wilson 2006). These complications are dealt with differently in each of the recently developed approaches above, resulting in a number of viable OT accounts of partial harmony cases such as Pulaar.

To illustrate how the more contemporary theories deal with these problems the following three approaches will be discussed below; Span theory; ABC harmony & Spread theory.

2.4.1. Span Theory

In span theory directionality is accounted for by two constraints first the Span Head right/left [ATR] (McCarthy 2004: 12) constraint which requires that the head segment of a span is either final or initial in that span depending on the directionality of the pattern (Sasa 2009). The next constraint contributing to directionality is Head Faith [ATR] (Sasa 2008). This is a faithfulness constraint which states that the head segment in the output is identical to that of the input in terms of its value for the feature [ATR]. Although these constraints are successful in selecting the optimal candidate whether they reflect a pattern of total or partial harmony, they require that [ATR] be defined as a privative feature. It is therefore pre-determined that [+ATR] is the marked value and active feature, which is vital to the assimilation process as span theory dictates that only privative features may spread.

(6) Example from Pulaar, word medial trigger (Sasa 2009: 103):

/hɛl-ir-dɛ/	Head-R	Head Faith [ATR]	S-Parse [ATR]	*Unparsed	Id [ATR]
a. (hel-i̲r)-dɛ ☺			*	*	*
b. hɛl(-i̲r-de)	*!		*	*	*
c. hɛl(-i̲r)-dɛ			*	**!	*
d. (hel-ir-de̲)		*!			**

In the tableau above candidate (6b) loses by violating the directionality constraint Head-R. Candidate (6d) loses by the Head Faith constraint, in which case the head segment in the output does not correspond with that in the input. This also relates to the ultimate direction of assimilation as it prefers the already [+ATR] segment [i] as the head. Candidate (6c) reflects a situation of no assimilation and therefore no harmony, however this candidate loses on the markedness constraint *unparsed. This constraint therefore prefers that more segments share the same feature value, forcing all segments into a span by placing violations on unparsed segments. Hence, candidate (6a) surfaces as the optimal candidate, reflecting a case of partial harmony in which spreading is regressive. It is clear that span theory is capable of providing a comprehensive account of complex harmony systems such as Pulaar. Both the issue of directionality as well as sour grapes is accounted for when [ATR] is considered privative.

2.4.2. ABC Harmony

The application of ABC harmony to the Pulaar case results in the following ranking: *ɪ/ʊ, *ǽ, Proximity, Ident I-O [ATR] Word-Final, Ident VV [ATR], Corr [-Hi], Corr [-lo], Corr V-V, Ident [ATR]. The markedness constraints eliminate all candidates comprised of vowels that do not exist in the inventory (Archangeli & Pulleyblank 1994, Bakovic 2000, Kràmer 2001). The proximity constraint limits the locality of the segments in correspondence, thereby accounting for cases of opacity (Rose & Walker 2004). The positional faithfulness constraint works to retain the [ATR] value of the trigger, which is the right most final vowel in the regressive harmony pattern (Petrova 2006, Kràmer 2001). The remaining faithfulness constraints include Ident VV [ATR], which requires that all segments in correspondence be identical for the feature ATR, and finally Ident [ATR] states that correspondent segments in the input and output must share the same value for ATR (McCarthy & Prince 1995, Walker 2009). The correspondent constraints include Corr [-Hi] which requires that all non-high segments be in correspondence, Corr [-lo] requires that all non-low segments correspond, and finally Corr V-V Requires that all vowels be in correspondence (Rose & Walker 2004, Walker 2009). This constraint ranking enables the ABC harmony approach to deal with unattested candidates in Pulaar and produce the correct directionality by assigning violations to a lack of correspondence. The sour grapes issue is also accounted for mainly as a result of the constraint on proximity thereby favouring adjacent correspondence and partial harmony.

(7) ABC Harmony tableau: example from Pulaar (Harmony with a blocker) (Sasa 2009)

/hɛl-ir-dɛ/	Id (fin)	Prox	Id VV [ATR]	Corr [-hi]	Corr [-lo]	Corr V-V	Id [ATR]
a. helirdɛ ☺				*		*	*
b. hɛlirde	*!			*		*	*
c. hɛlirdɛ		*!	**				

d. helirde	*!				*	*	**

The sub-optimal candidates (7b) & (7d) are eliminated by the first constraint which prohibits the [ATR] alternation of the word final vowel. The final vowel in the attested form must therefore appear in the retracted form on the surface. Hence, the partial harmony candidate (7c) is eliminated next, by the constraint on proximity. The partial harmony reflected in this candidate is a case of transparency, in which two of the three segments share the feature [-ATR]. The proximity constraint ensures that agreement is prioritised between adjacent segments. Hence, the attested form candidate (7a) is considered optimal, in this form the final vowel remains faithful for [ATR] and the preceding adjacent vowels agree for [+ATR].

2.4.3. Spread Theory

Finally, the original spread approach requires that every segment be linked to one feature (Padgett 1997). However, this constraint will always select the unattested form in harmony patterns which demonstrate directionality such as the examples from Pulaar. Again, there is the sour grapes issue in which case the constraint ranking favours the candidate with no harmony rather than one with partial harmony (Wilson 2006). However, the modified constraint Spread [F] –Left (Sasa 2006) inserts a directionality requirement on the original spread constraint. This modification to the spread constraint enables the approach to accurately determine the attested candidate for Pulaar harmony, even in cases which require partial harmony agreement. It is therefore concluded in the literature, that the development of the spread constraint to include a specification for directionality provides a solution to both empirical issues. By placing a directional specification on the spread constraint the sour grapes problem is also resolved, as the constraint ranking now prefers candidates which exhibit regressive spreading and partial harmony above those which display no harmony. The spread constraint ranking presented for Pulaar (Sasa 2009) includes the following constraints; Spread [ATR]-L (Sasa 2006), Id [ATR](Fin) (Petrova 2000, Krämer 2001), Id [ATR](rt) (Beckman 1997), Ident [ATR] (McCarthy & Prince 1995), *[+Hi,-ATR], *[+Lo,+ATR] (Archangeli & Pulleyblank 1994, Bakovic 2000, Krämer 2001). As discussed above the modified spread constraint requires that only segments to the left of the trigger share the same value for the feature [ATR]. There are three faithfulness constraints the first, Id [ATR] (Fin) contributes to achieving directionality by ensuring the trigger retains its value for [ATR]. Id [ATR] (rt) then retains the [ATR] specification of segments in the root, and the final faithfulness constraint simply places a general restriction on the alternation of values for [ATR].

(8) Spread Harmony tableau: example from Pulaar (Harmony with a blocker) (Sasa 2009)

/hɛl-ir-dɛ/	Id [ATR] (Fin)	Spread [ATR]-L	Id (Root) [ATR]	Id [ATR]
a. heli̲rdɛ ☺				*
b. hɛli̲rde	*!	*		*
c. hɛli̲rdɛ		*!		
d. helirde̲	*!			**

The first constraint, which functions to retain the [ATR] value of the word final vowel is violated by candidates (8b) & (8d). As in the case of the ABC harmony tableau above, the correct default form of the final mid-vowel is ensured by the positional faithfulness constraint. The deciding constraint however, is the spread constraint with the leftward directionality requirement. This constraint does not enforce a particular trigger, but eliminates the final sub-optimal candidate by requiring harmony between the segments which are not restrained by positional faithfulness.

2.4.4. Discussion and comparison

In both the tableaux presenting ABC harmony as well as the one above presenting harmony through spread, the unattested candidates are eliminated by violations on at least one faithfulness constraint retaining the underlying specification for the feature [ATR], as well as a constraint tailored to produce the attested directionality. In the case of the ABC approach this includes what is referred to as directionality by lack of correspondence, which is produced by constraints Corr [-Hi] & Corr [-lo]. The spread approach relies on the modified constraint Spread [ATR]-L to eliminate candidates with the unattested directionality. These approaches are therefore not entirely equip to handle the empirical issues explicated above. Each make use of the same basic ranking structure in order to achieve the correct harmony pattern. The highest ranked constraints in each case account for the limitations of the vowel inventory, after which the harmony pattern and directionality are accounted for through some form of spread or correspondence constraint paired with a dependence on faithfulness to some underlying form or feature value. These approaches are therefore successful in dealing with a complex harmony pattern, with the prerequisite of various postulations in order to express the capacity of each approach. Span theory requires that segments be parsed into spans in order for the constraints to function effectively. Harmony with ABC is dependent on privativity, and the postulation that only one feature value is marked. The biggest challenge to the uniformity of these approaches however is the reliance on an underlying form. The need for segments of a candidate to remain faithful to a predetermined from in order to achieve the necessary directionality.

(9) Comparison of constraint ranking structures

Span Theory	ABC Harmony	Spread Harmony
*ɪ/ʊ, *ǽ Head[ATR]-R Head Faith S-parse[ATR] Ident [Hi] *e/o Ident[ATR]	Prox *ɪ/ʊ, *ǽ Ident(Fin) Ident VV[ATR] Corr[-Hi],Corr[-Lo],CorrV-V Ident I-O [ATR]	*ɪ/ʊ, *ǽ Ident [ATR](Fin) Spread [ATR]-L Ident (root) Ident [ATR]

3. Data and Generalisations

3.1. Introduction

As is the case with many African languages of the Bantu family a tongue root distinction between vowels is a recorded occurrence in isiXhosa, and has therefore been phonetically analysed and described. The feature Advanced tongue (ATR) root has been reported to pattern between vowels within the prosodic word, resulting in a system of featural harmony between different vowel segments. The ATR harmony pattern found in isiXhosa is quite complex, the pattern may occur as either transparent or opaque. Harmony cases therefore vary between favouring lexical items bearing more segments with the same value for ATR, or instead lexical items which have adjacent segments sharing the same feature value for ATR. Depending on the phonology of each individual lexical item, the attested form may represent a case of total harmony, partial harmony or no harmony. The outcome or attested form is therefore not always uniform, which must be accounted for by the generalisations used to describe it. The idiosyncrasies observed in the recorded pattern call for generalisations specifically relating to the function of different vowels, as well as generalisations describing the structural complexities of the pattern. Although the mid-vowels in isiXhosa are the only vowels in the inventory with a harmonic counterpart for the feature ATR, all vowels contribute to the resulting harmony pattern.

3.1.1. Generalisations

As stated in chapter one, all examples follow the same transcription convention in which vowels are presented in IPA format, and all consonants are presented in the orthographic form. The generalisations used to define the harmonic pattern can be categorised into two sets. Those which describe the function of each vowel within the harmony system, and those which describe the more general structure of the system. The initial set of bulleted generalisations describe the function of the different sets of vowels.

(1) Summary of generalisations: vowel functions

- **Mid-vowels are the only vowels that alternate for the feature ATR**
- **High vowels are the only vowels with the capacity to trigger vowel raising**
- **The low vowel [a] acts as an opaque segment when positioned between a preceding mid-vowel and a high vowel**

The structure of the harmony pattern then include specifics relating to the active feature value [+ATR], the latter is therefore considered the default option. Also included in the pattern are specificities concerning directionality, specifically the direction of [+ATR] feature spreading, and finally the extent to which this rule may be applied.

(2) Summary of generalisations: system structure

- **[+ATR] is the active feature in the harmony pattern of isiXhosa**
- **All spreading of the feature [+ATR] is regressive**
- **The rule of [+ATR] spreading is iterative**

3.2. Morphological structure of words: isiXhosa

Both verbs and nouns were included in the collected data, as the isiXhosa morphology allows for numerous affixes to be attached to either of these grammatical classes. Although nouns are not deconstructed in exactly the same manner as verbs, nouns for example do not bare a detachable final vowel. However, including both grammatical classes has enabled the inclusion and processing of a wider variety of data. The surface constructions produced for the analysis of this harmony system were required to contain a particular sequence of vowels in order to test a specific generalisation. For example, constructions with [a] positioned between a mid-vowel and a high vowel are necessary to test the opacity of the low vowel, and furthermore to answer the question of harmony blocking. Constructions which have a high vowel positioned between two mid-vowels are necessary in order to test directionality. These constructions would indicate whether spreading affects only the preceding mid-vowel or both mid-vowels preceding and following the high vowel. To test the existence of a default option, constructions which had roots with only mid-vowels are used. The understanding is that the default realisation of the mid-vowel is revealed by eliminating the influence of both high and low vowels. At this level the collected data was selected and constructed in order to test each of the findings presented in Jokweni & Thipa (1996).

The recorded constructions were selected according to the surface vowel combinations produced. These phonological combinations are produced by means of the agglutinating nature of the isiXhosa morphology. There a number of ways in which one may choose to deconstruct a prosodic word in isiXhosa, some being more in-depth than others (Doke 1954). For the purposes of this analysis however, a morphological deconstruction is considered unimportant as it has yet to contribute to an understanding of the harmony process. Hence, the words have been morphologically deconstructed into Prefix- Root/verbal base-suffix/final vowel depending on the grammatical class to which they belong (Poletto 1999; Schadeberg 2003).

(3) Verb structure

[ɛbhalekiswa]
ebhalekiswa

e-	bhalek	-is	-w	-a
prog	run	caus	pass	fv
prefixes	root	suffixes		final vowel

'(s)he is being caused to run'

[ɛbɛbhadlile]
ebebhadlile

ebe-	bhadl	-ile
prog	angry	past
prefixes	root	suffixes

'(s)he was caused to become angry

(4) Noun structure

[umsebenzi]
umsebenzi

um-	sebenz	-i
1a.sg	work	agent nom
prefix	root	suffix

'worker'

[oorhulumente]
oorhulumente

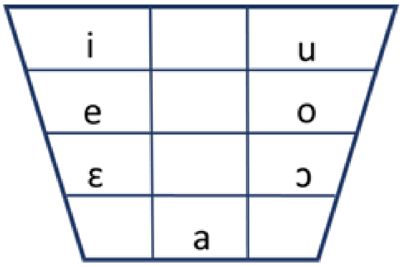
oo-	rhulumente
1a.pl	government
prefix	root

'governments'

In isiXhosa the archetypal noun and verb forms consist of a stem, this is the base of the word which may be described as the primary lexical component within a word as it carries the most salient semantic information. Preceding the base, the noun phrases also consist of a nominal prefix such as the prefix [o-] in [orhulumɛntɛ] above. The nominal prefix will serve a finite number of functions, including the class to which a particular word belongs e.g. class 1a [u-] – [urhulumɛntɛ]. The nominal prefix is also used to indicate number, whether the object referred to is singular or plural. Hence, in the example above the class prefix [u-] becomes [o-] to indicate a reference to more than one government. The morphology of verbs in isiXhosa consist of a wider variety of attachable affixes, which perform a range of functions (Tshabe, Shoba, Mini, Pahl, Pienaar & Ndungane 2006). The attachable prefixes include inflectional morphemes used for aspect, tense, mood, or negation, as well as reference to a particular noun class or person etc. Unlike nouns, verb stems can be further deconstructed to include the verbal base as well as a final suffix or vowel. This kind of deconstruction is also possible in nominal cases in which the noun has been derived from a verb, however in most cases as in the example above, the final vowel forms part of the nominal base which cannot be further morphologically deconstructed (Schadeberg 2003).

3.3. Generalisations

(5) Vowels of Xhosa (from Jordan 1966 & Pahl 1967)



isiXhosa has a seven surface vowel system with two sets of mid-vowels (Jokweni & Thipa 1996, Jordan 1966, Pahl 1967, Tshabe, Shoba, Mini, Pahl, Pienaar & Ndungane 2006). According to Jokweni & Thipa (1996) these vowels are distinguished by their value for the feature advanced tongue root or [ATR]. The feature values [+-ATR] are non-contrastive in nature and are closely related to vowel height. [+high] = [+ATR], [-high] = [-ATR] (Jokweni & Thipa, 1996). The feature distinguishing the mid-vowels is recognised as ATR, while the alternate forms are also referred to as semi-open [-ATR] and semi-closed [+ATR] vowels by Jokweni & Thipa (1996) as well as Pahl (1967) & Sibanda (2008). The tongue root feature itself is still indefinite in terms of its place in featural geometry, and whether or not it is simply an offshot of the value of the height node (Wiswall 1991). ATR has also been proposed as a separate node dominating the unary feature [ATR] beneath the place node. However, it has also been suggested that this feature should be dominated by the articulatory node (Wiswall 1991).

(6) Vowel feature tableau (adapted from Jokweni & Thipa 1996)

	i	e	ɛ	a	o	ɔ	u
high	+	-	-	-	-	-	+
low	-	-	-	+	-	-	-
ATR	+	+	-	-	+	-	+

Discussion is therefore in progress in order to determine whether or not there is enough evidence to advocate for the occurrence of ATR as a dominating feature, or whether it is simply a consequence of the height value of a vowel. The tongue root feature recorded cross-linguistically in the vowel harmony of African languages has been described in a number of ways, referred to as breathy, tense and raised height (Fulop, Kari & Ladefoged 1998). In most cases substantial movement of the tongue root has been recorded through the use of radiographic data. The consensus is therefore that in adopting this feature the speaker is manipulating the size of the pharynx in order to change the quality of the vowel (Lindau 1978). It is widely acknowledged however that the disparity in pharynx size between the vowels which harmonize differently can be recorded and identified acoustically. Noticeable by the value of the first and second formants it is also clear why linguists previously described this type of harmony as a height difference between vowels (Lindau 1978). It is also suggested that this feature would more accurately be referred to as expanded, due to the expansion of the pharynx (Halle & Stevens 1969). Referring to this feature as expanded as oppose to ATR is thought to be accurate as it takes into account more than the necessary articulation.

3.4. Vowel participation in the harmony process

3.4.1. Mid-vowels alternate for the feature ATR

Mid-vowels are the segments which undergo harmony, hence orthographic /e/ & /o/ are the only vowels which alternate between harmonic counterparts with different feature values for [ATR]. In (7&8) below there are two representations of the same lexical item. The first includes two mid-vowels sharing the feature value [+ATR] with the attachment of the causative morpheme /isa/. In this case the mid-vowels are followed by the high vowel /i/ which triggers the advancement of the two preceding mid-vowels. The second form of this lexical item is the basic form with no morphological alteration. In this case the two mid-vowels share the feature value [-ATR]. Without the attachment of the causative suffix there is no following high vowel to trigger vowel raising, hence the mid-vowels occur in the retracted form. (9) Below is a representation in the past tense form, however even with the attachment of the past tense morpheme the mid-vowels still occur in the retracted form. The explanation for the lack of vowel raising in this case is blocking, which will be discussed later. The function of these examples is simply to exhibit mid-vowel alternation for the feature ATR. These examples as well as those to follow are all presented using orthographic script, with only the vowels presented in IPA.

(7) /e/+/bon/+/isa/ → [e-bon-isa] ‘him/her/it showing’

(8) /e/+/bon/+/a/ → [ɛ-bɔn-a] ‘him/her/it see

(9) /ebe/+/bhadl/+/ile/ → [ɛbɛ-bhadl-ilɛ] ‘he was caught off-guard’

3.4.2. High vowels trigger vowel raising

The advanced mid-variants [e] & [o] occur only when followed by a high vowel, this has been observed and reported on in the findings of previous literature Jokweni & Thipa (1996) in the form of a [+ATR] trigger requirement, as well as in the novel data presented here. The presence of the high vowel has an effect on all preceding mid-vowels, causing them to occur in the advanced form. These forms do not appear in any other phonological context and are therefore concluded to be the result of the raising effect produced by the following high vowel. In (10) a) below all four of the front mid-vowels preceding the high vowel [i] appear in the advanced form, all of these vowels have therefore been raised. Example (10) b) exhibits a case in which only one of the mid-vowels in the word has been raised. The second mid-vowel is not raised as it is the final vowel in the word and therefore does not have a following high vowel trigger. Examples (10) c-j) are additional instances of the advanced mid-variant preceding a high vowel. The other end of this raising phenomenon indicates that without a following high vowel the mid-variant will occur in the retracted form. This is demonstrated in examples (11) a-e) in which none of the mid-variants precede a high vowel and therefore all occur in the retracted form.

(10) Mid-variants preceding a high vowel [e/o-i/u]:

a. /e/+/ntlekeleni/ → [e-ntlekeleni] ‘in a calamity’
b. /ndi/+/on/+/ile/ → [nd-on-ilɛ] ‘I did wrong’
c. /i/+/bhedengu/ → [i-bedengu] ‘liar, thief’
d. /e/+/sikolweni/ → [e-sikolweni] ‘at school’
e. /i/+/xelegu/ → [i-xelegu] ‘dirty person’
f. /ndi/+/feket/+/ile/ → [ndi-feket-ilɛ] ‘I played’
g. /ndi/+/sebenz/+/ile/ → [ndi-sebenz-ilɛ] ‘I worked’
h. /komzi/ → [komzi] ‘feeling of community’
i. /u/+/bhebhetyu/ → [u-bhebhetyu] ‘ugly person or thing’
j. /u/+/tatomkhulu/ → [u-tatomkhulu] ‘grandfather’

(11) Mid-variants not preceding a high vowel [ɛ/ɔ]:

a. /ndi/+/on/+/e/ → [nd-ɔn-ɛ] ‘I did wrong’

b. /i/+/rekene/ → [i-rɛkɛnɛ] ‘a boy’s catapult sling’

c. /i/+/qokobhe/ → [i-qɔkɔbhɛ] ‘empty rigid container or shell’

d. /i/+/sikolo/ → [i-sikɔlɔ] ‘school’

e. /i/+/qengele/ → [i-qɛngɛlɛ] ‘a swollen part of the body e.g. tumor’

3.4.3. Low vowel causes blocking

The low vowel [a] does not possess a harmonic counterpart distinguished by the feature ATR and is therefore described as neutral (Van Der Hulst & Van De Weijer 1995). Although neutral in nature the function of this vowel results in cases of opacity. The insertion of [a] between the high vowel and a preceding mid-vowel results in the blocking of [+ATR] spreading. Any mid-vowels first preceding [a] followed by a high vowel trigger, will not undergo the effects of vowel raising. Example (12) a) below should have one advanced mid-vowel as the vowel [o] occurs preceding the high vowel [i], however because of the interference of the low vowel the mid-vowel occurs in the retracted form [ɔ]. In (12) b) below the prefix [oo-] should occur in the advanced form according to the rule of high vowel raising, however due to the insertion of a low vowel between the high vowel and preceding mid-vowel prefix, this process of assimilation does not take place. There are however instances in which some degree of raising is permitted before any interference by the low vowel. Hence, in example (12) d) there is one mid-vowel which occurs between a preceding low vowel and a following high vowel, this is the only advanced mid-variant in the construction. The high vowel has therefore caused [+ATR] assimilation on the preceding mid-vowel before the interference of the low vowel. Raising is then blocked by the low vowel prohibiting the spread of [+ATR] to the next set of preceding mid-vowels. In each of the cases (13) a-e) the mid-variant directly preceding the high vowel appears in the advanced form as it occurs before the low vowel blocker. Any further spreading of [+ATR] is then prohibited due to the opacity of the inserted low vowel.

(12) Mid-variants preceding a low vowel, followed by a high vowel [ɛ/ɔ-a-i/u]:

a. /kondlaliw/+/e/ → [kɔndlaliw-ɛ] ‘the bed is ready’

b. /oo/+/malum/+/e/ → [ɔː-malum-ɛ] ‘uncles’

c. /ebe/+/bhadl/+/ile/ → [ɛbɛ-bhadl-ilɛ] ‘he was caused to become angry’

d. /ebe/+/qabukel/+/isa/ → [ɛbɛ-qabukel-isa] ‘he was caused to look back’

e. /u/+/lwandlekazi/ → [u-lwandlɛkazi] ‘ocean’

(13) Mid-variants preceding and following a low vowel & preceding a high vowel [ɛ/ɔ-a-ɛ/ɔ-i]:

a. /ebe/+/bhadek/+/ile/ → [ɛbɛ-bhadek-ilɛ] ‘he was caught off-guard’

b. /ebe/+/qabel/+/isa/ → [ɛbɛ-qabel-isa] ‘cause to ascend or go over the top’

c. /e/+/bhakontini/ → [ɛ-bhakontini] ‘outdoor brick made oven’

d. /e/+/sandleni/ → [ɛ-sandleni] ‘applause’

e. /u/+/nesandlozi/ → [u-nɛsandlozi] ‘wide starring eyes’

3.5. Structural generalisations

The initial generalisations presented above reflect the participation of the different vowels within the harmony system, we now shift to generalisations which describe the structural process independent of the roles of individual vowels.

3.5.1. Mid-vowels default to [-ATR]

ATR functions as a singular feature of which one value is active and the other serves as a default option. As established in the findings of Jokweni & Thipa (1996) the [+ATR] mid-variants occur only when triggered by a following high vowel, while the [–ATR] variants occur in all other contexts. The ideal contexts for testing this default effect include combinations in which mid-vowels occur with neither a following high nor low vowel. The following examples display lexical items with only mid-vowels, and these mid-vowel occurring in the retracted form.

(14) Mid-vowels without any following high/low vowel influence [ɛ/ɔ]:

a. /e/+/mhloph/ → [ɛ-mhlɔpɛ] ‘white’

b. /bhetele/ → [bhɛtɛlɛ] ‘better’

c. /u/+/qhoqhoqho/ → [u-qhɔqhɔqhɔ] ‘windpipe/trachea’

d. /u/+/rhwebo/ → [u-rhwɛbɔ] ‘commerce/trade/trading’

e. /u/+/msebe/ → [u-msɛbɛ] ‘eyelash’

f. /u/+/qwethe/qwethe/ → [u-qwɛthɛ-qwɛthɛ] ‘irritation in the mouth or throat’

g. /i/+/thole/ → [i-thɔlɛ] ‘calf’

h. /i/+/sondo/ → [i-sɔndɔ] ‘corner or edge of a blanket’

i. /i/+/qongwe/ → [i-qɔngwɛ] ‘periwinkle (sea snail)’

j. /i/+/rekethe/ → [i-rɛkɛthɛ] ‘racket, racquet’

3.5.2. Directionality of spreading

The process of active feature spreading can now be amended to include the notion that spreading is always regressive. Therefore, the spreading of the feature [+ATR] is only introduced to a mid-vowel which occurs preceding the high vowel trigger. In the first example below the two mid-vowels preceding the high vowel occur in the advanced form while the final mid-vowel is retracted. It might be argued that the initial high vowel is the cause of the advancement of the following mid-vowels, however the second example displays a case of a high vowel followed by two mid-vowels which both occur in the retracted form. The conclusion is that [+ATR] spreading is bound both by an appropriate trigger, as well as the position of this trigger in relation to the alternating mid-vowels. Vowel raising may only occur when the high vowel is the right-most segment and can therefore effect any preceding mid-vowels. The system is one of regressive spreading.

(15) Mid-variants preceding and following a high vowel [e/o-i/u-ɛ/ɔ]:

a. /i/+/monez/+/ile/ → [i-monez-ilɛ]	'that thing satisfied him'
b. /e/+/sile/ → [e-silɛ]	'donkey'
c. /e/+/songiweyo/ → [e-songiwɛyɔ]	'savings'
d. /i/+/bhekile/ → [i-bhekilɛ]	'tin/can'
e. /i/+/qobiso/ → [i-qobisɔ]	'threat or pretence of harm'
f. /i/+/saneliso/ → [i-sanelisɔ]	'anything that satisfies, fills, pleases'
g. /ndi/+/qond/+/ile/ → [ndi-qond-ilɛ]	'I understood'
h. /oo/+/loliwe/ → [o:-loliwɛ]	'trains'
i. /u/+/bhelushe/ → [u-bhelushɛ]	'yellow fever'
j. /u/+/swelek/+/ile/ → [u-swelek-ilɛ]	'he/she died'

3.5.3. Locality generalisation

The final generalisation accounts for this harmony process spreading to unbounded segments. The process of high vowel raising is therefore not constricted by any morphological domain. This is illustrated by examples such as (16) a-j) in which mid-variants alternate within the stem as well as the prefixes. Spreading of the feature [+ATR] onto the mid-vowel is not confined to a morphological domain such as the prefix, suffix or root. Therefore, unlike many African languages the harmony system is not stem-controlled (Van der Hulst & Van de Weijer 1995). In isiXhosa it is clear that no particular morphological domain determines whether the construction will be dominated by [+ATR] or [-ATR] harmony. Recorded data exhibit all morphemes undergoing alternations for ATR, including the stem. Although not constricted by a morphological domain, the syllabic or phonological context does influence the raising process. The insertion of a particular segment within the phonological context will determine whether [+ATR] spreading is initiated, and whether it will be permitted to extend across the rest of the phonological domain. Examples (17) a-e) display instances in which mid-vowels within the root have been raised due to a following high vowel trigger in the suffix. [+ATR] spreading is then not permitted to progress to the mid-vowels in the prefix due to the insertion of the low vowel. Hence, the reason the mid-variants occur in the retracted form in the prefix of example (17) a), is solely as a result of interference by the low vowel segment. The phonological context therefore prohibits the extension of the [+ATR] spread rule to the word initial mid-vowels. The effects of this rule are therefore generally iterative and will continue to spread across any preceding mid-vowel undergoers, unless debilitated due to the interference of a particular segment.

(16) Raising permitted to unbounded segments:

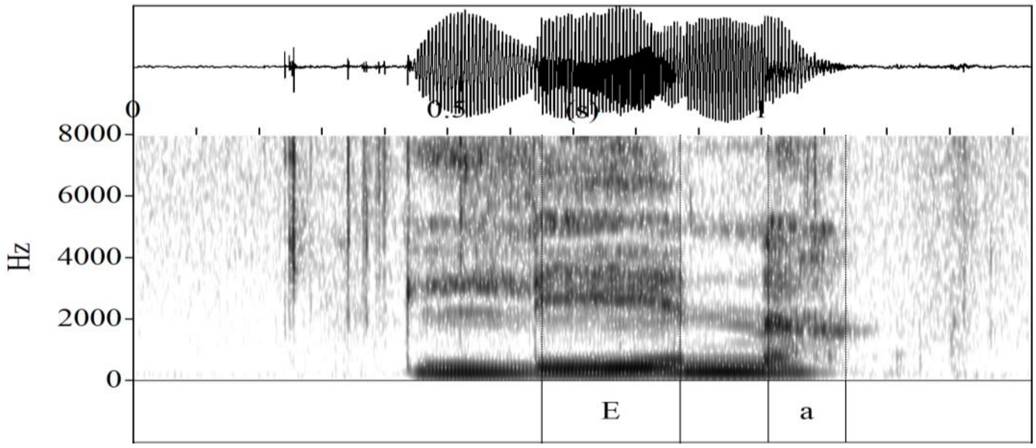
a. /e/+/silonden/+/i/ → [e-silonden-i]	'hard covering crust e.g. of bread'
b. /o/+/kwemilenze/ → [o-kwemilɛnzɛ]	'state of legs: to be drawn up'
c. /oo/+/thixo/ → [o:-thixɔ]	'Gods'
d. /ne/+/e/+/sile/ → [ne-e-silɛ]	'with the donkey'
e. /o/+/renji/ → [o-renji]	'orange'
f. /no/+/muntu/ → [no-muntu]	'someone'
g. /nd/+/enz/+/ile/ → [nd-enz-ilɛ]	'I made'
h. /ebe/+/phengulul/+/a/ → [ebe-phengulul-a]	'he/she turn upside down'
i. /ebe/+/bheculul/+/a/ → [ebe-bhecululua]	'he/she opened their eyelid'
j. /e/+/sekenz/+/iswa/ → [e-sekenz-iswa]	'he/she is caused to work'

(17) Raising blocked by phonological context:

a. /ebe/+/qabel/+/isa/ → [ɛbɛ-qabel-isa]	'caused to ascend or go to the top of'
b. /ebe/+/bhadek/+/ile/ → [ɛbɛ-bhadek-ilɛ]	'he was caught off-guard'
c. /ebe/+/yalek/+/isa/ → [ebe-yalek-isa]	'to teach one lesson'
d. /e/+/sandleni/ → [ɛ-sandleni]	'applause'
e. /ebe/+/qabukel/+/isa] → [ɛbɛ-qabukel-isa]	'to look back'

3.6. Additional observations:

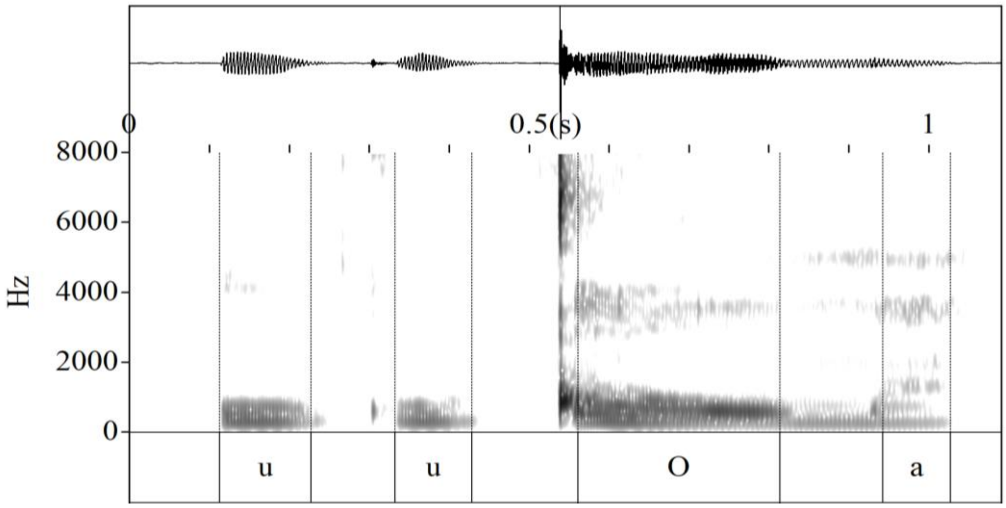
In the process of the acoustic analysis certain observations were made which are not dealt with in this paper. In some cases, the quality of mid-vowels is determined by factors other than harmony. These are not analysed in depth in this thesis, but are noted here. As seen in the vowel charts presented in the acoustic analysis chapter there have generally been tokens form each segment which are not representative of the predicted acoustic composition of that segment. The tokens which form this band of outliers is comprised of particular segments which have been annotated correctly according to their phonological context, however due to additional phonetic influence they are not truly representative of the ATR value for which they have been annotated. The bulk of these misrepresentative tokens have been annotated as [-ATR] and should occur in the retracted form due to their phonological context. However, presented in figures 18 & 19 below the retracted mid-vowel preceding a nasal consonant share F1 values similar to the [+ATR] vowels as opposed to those which are considered [-ATR] such as the low vowel.



(18) Figure: lena ‘This (one) thing’

The mid-front vowel [e] in the word ‘lena’ is annotated as the retracted form [ɛ] in the acoustic representation above. This annotation is aligned with the generalisations as there is no high vowel trigger in the relevant construction and therefore the vowel should occur in the retracted default form. However, as displayed in figure (18) above the F1 resonant value of this segment is 362.4 Hz. This value is well above the average F1 value of the retracted default variants calculated in the acoustic analysis section, and is in fact greater than the average F1 value of the high vowels.

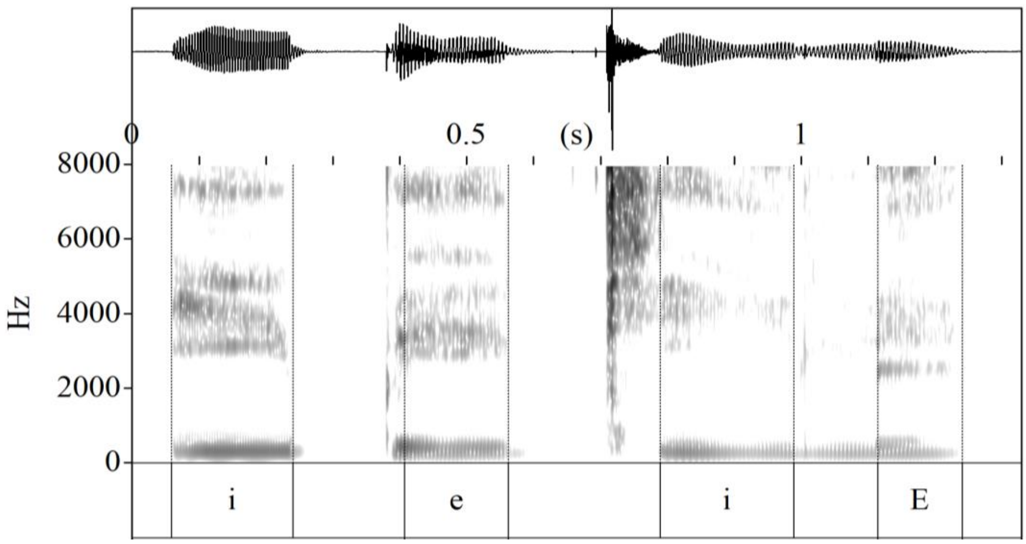
[ɛ] F1	[ɛ] default average F1	[i] average F1	[u] average F1
362.4 Hz	508,996	340,617	359,952



(19) Figure: ukuqoma 'feel pain in the mind or heart'

The back mid-vowel in figure (19) above is also annotated according to the predicted ATR value, which is [-ATR] as there is no high vowel trigger to initiate raising. However, this mid-variant also shares a closer resemblance to the F1 value of a [+ATR] vowel rather than that expected of a retracted form. The presence of the nasal consonants are therefore causing any preceding vowels to occur as 'closed' rather than 'open' variants.

[ɔ] F1	[ɔ] default average F1	[i] average F1	[u] average F1
422.7	498,674	340,617	359,952



(20) Figure: ibhekile 'tin/can'

The same effect is seen in the retracted mid-variants following the lateral approximant [l]. All word final mid-variants should occur in the retracted form as there are no following high vowel segments to produce an alternation. Therefore, all mid-variants in the word final position has been annotated as [-ATR]. However, as demonstrated in the acoustic representation above the final mid-variant shares a similar F1 value to the raised mid-variant preceding the high vowel. Certain instances in which the final mid-vowel is preceded by [l] result in the final vowel appearing more 'closed' therefore more advanced. This consistent influence on the final vowel may be the reason for the insignificant discrepancy between the mid-advanced and retracted variants extracted from this particular phonological context.

As discussed below, extensive provisions are taken to avoid any consonantal interference on the vowel segment. However, as demonstrated in the figures above there are two sets of consonants in isiXhosa which have a strong influence on either the preceding or following vowel. Similar co-articulatory results have been reported to occur in other Nguni languages such as Zulu (Doke 1954, Harris 1987). Syllabic [m] causes more closure on any preceding vowel regardless of the ATR value of the vowel in the subsequent syllable. In isiXhosa however it appears this form of closure may be the influence of any nasal consonant. The closure caused by the lateral approximant [l] may be a result of vowel like nature of approximants. Hence, there is no real closure or end to the articulation of the initial vowel and this may therefore cause some kind of carry over articulatory effect (Gay 1977).

3.7. Collected data

The data required to investigate the findings above include specific combinations of vowels and morphemes. Each combination will have the high and low vowel within particular proximity and physical relation to the mid-vowel. The examples above reflect different vowel combinations relevant to each generalisation and the required phonetic context. As each of the findings refer to harmony effects on the mid-vowel in different phonetic contexts, the data used to test each of these must position the mid-vowel in a specific phonological context. The collected data described in conjunction with the generalisations above has therefore been divided and analysed to avoid the interference of diverse variables. The agglutinative nature of the language provides an accessible and uniform approach for collecting the necessary data. The array of attachable affixes ensure evidence of all mid-vowel alternations as well as the cases in which raising does not occur. The grammatical classes included in this study are limited to verbs and nouns. As these classes do not accept the same morphological modifications (Doke 1954) the affixes attached to each are presented below.

3.7.1. Constructions under investigation

As mentioned above an understanding of the isiXhosa morphology is necessary in order to produce words with the required vowel sequences. Hence, there are a number of prefixes and suffixes which were most useful in eliciting the relevant phonetic contexts. Using the prefixes below the height of vowels preceding and following any mid-vowel were manipulated to demonstrate any harmony affects as a result of their positioning. The following list of morphemes is sufficient for the investigation of ATR harmony. The addition of these morphemes onto different roots offers a series of phonological contexts. Hence, with the mid-variant prefixes for example one is able to test the limitations of raising. The high vowel suffixes are added to produce instances of raising etc.

(21) Verbs:

Prefixes:		Suffixes:	
e-	2nd person pronoun	-e	past tense
ebe-	subject marker & tense marker	-ile	perfect
w-	subject marker	-is	causative
nd-	1st person pronoun	-w	passive
u-	3rd person pronoun	-a	Final vowel

(22) Nouns:

Prefixes:	
u-	Class 1a
oo-	plural
i-	

3.7.2. Constructions categorised to test generalisations

The prefixes which include a mid-vowel or attach to a mid-vowel-initial root, are placed in categories which test the extent of high vowel spreading. Testing the extent of spreading across morphological items demonstrates both the iterative nature of the process, as well as the domain in which harmony occurs. Prefixes which include high vowels are categorised as constructions testing high vowel spreading and directionality. Constructions in which the high vowel resides in the prefix indicate whether a preceding high vowel may trigger raising or affect the following mid-vowel by means of co-articulation. Having constructions in which the high vowel precedes the mid-vowel also provides the opportunity to compare the effect of a high vowel trigger occurring to the right of the mid-vowel which would equal regressive spreading, versus the effect of a preceding high vowel. Hence, without attachable prefixes the constructions available to test such generalisations would be limited. As presented in the tableau above, the most commonly used suffixes include the two past tense inflectional suffixes, the first of which includes only a mid-vowel, and again allows for the testing of the default form, domain and directionality. The second past tense suffix is used mainly to test directionality, as it may be attached to a root containing a final mid-vowel, thereby ensuring a sequence of mid-vowels both preceding and following a high vowel. The final suffix is the causative which is easily attachable to various verbal roots. With regard to the constructions included in the collected data the causative suffix is only attached to verb roots which contain mid-vowels. These constructions function as an indication of any acoustic effect on the mid-vowel caused by the following high vowel.

3.8. Appendix: list of words considered in data set

The list of words produced below is arranged according to the particular vowel sequences they share, and consequently according to the harmony patterns expected to occur within each phonological context.

[E/O]

afikane	Scented grass, used to make necklaces
ubawo	Father
hememe	Ideophone (sense: you’re in trouble)
impendulo	Answer
inthlekele	Calamity
inyaniso	Truth
iqekebhe	Unusually large, solid thing, usually circular
iqekele	A large head on a lean person
iqengele	A swollen part of the body e.g. tumor
iqokobhe	Empty rigid container or shell
iqongwe	Periwinkle (sea snail)
irekene	A boy’s catapult sling
irekethe	Racket, racquet
isakramente	Sacrament
isamente	Cement
isende	Testicles

ishweshwe	Concubine
isibalo	Maths sum
isikolo	School
isiqaqasholo	Something had and stiff
isiqubenge	A strong/powerful person or animal
isiqusheko	Bread
isondo	Corner or edge of a blanket
isongololo	Millipede
ithole	Calf
malume	Uncle
ndenze	I made
ndibaleke	I ran
ndifekete	I played
ndilese	I read
ndimkyelele	Visit him
ndimxlelele	I told
ndisebenze	I worked
ndithenge	I bought
ndithethe	I spoke
ndityelele	I visited
ndive	I heard
ndivile	I heard
ndone	I did wrong
ukuthiqeke	Stand apart or alone
ukuthiqelem	A bullet: to enter the body and remain there
ukuthiqheke	Split/cleave into two
ukuthiqhekre	Of day: break/dawn
ukuthiqhokro	Of trees: burst into blossom
ukuthiqwempe	Tear, break or cut
umbhidile	To be deceived
umbhido	Deception, fraud, trick
umendo	Highroad
umlilo	Fire
umqobo	Something long and thick
umsebe	Eyelash
uqengehe	Cheat/mean trickster
uqhomfo	Abortion
uqhoqhoqho	Windpipe/trachea
uqulathe ntoni emlonyeni	What have you in your mouth?
uqwethe qwethe	Irritation in the mouth or throat
urhwebo	Commerce/trade/trading
urongo	Wrong/incorrect
usendo	A cut of meat eaten at the initiation of girls
usipho	Proper noun
uthixo	God
uxulumente	Government
wavisiwe	They heard

[E/O…i/u]

akademiki	Academic
akijologi	Archaeology
amabhelubhentsu	Unexpected measures taken to meet an emergency
amendu	Speed
enthlekeleni	In a calamity
esikolweni	At school
etholeni	On the calf
ianatomi	Anatomy
iastronomi	Astronomy
ibaijaskopu	Cinema
ibhedengu	Liar, thief
ixelegu	Dirty person
ikhokeli	Leader
iqhoni	Patchwork/petticoat of various colours
isambreli	Umbrella
isayensi	Science
isiqwempu	A piece of cloth or cord
isisimpothi	Sympathy
isongiwu	
iveki	Week
komzi	Feeling of community
nomuntu	Someone
orenji	Orange
ubhebhetyu	Ugly person or thing
ukungapheleli	To serve more than one purpose
umananeli	Goods exchange
umqwebedu	Hard, dry, sterile ground
umrhwebi	A trader/businessman
utatomkhulu	Grandfather

[E/O…i…E/O]

babakleyile	Fight, quarrel
esile	Donkey
esilondeni	Hard covering crust e.g. of bread
esongiweyo	Savings
ianayestile	Anniversary
ibaiyisekile	Bicycle
ibhekile	Tin/can
ibhengile	Bangle
idwelibhekile	Look at it
imonezile	The thing that satisfied him
iqobiso	Threat or pretence of harm
iqwebiso	Act of threatening or pretending to do something
isaneliso	Anything that satisfies, fills, pleases

isiqalekiso	Curse
izelibhekile	The can is full
ndenzile	I made
ndibalekile	I ran
ndifeketile	I played
ndilesile	I read
ndimxlelelisisile	I visited
ndipendule	I answered
ndipendulile	--------------
ndiqondile	I understood
ndisebenzile	I worked
ndithengile	I bought
ndithethile	I spoke
ndiyenzile	I made
ndonile	I did wrong
neesile	With the donkey
okwemilenze	State of legs: to be drawn up
oololiwe	Trains
oothixo	Gods
ooxulumente	Governments
ubhecululo	Academic investigation, thorough search or assessment
ubhelushe	Yellow fever
ukukwalesile	To sleep
ukuthi bhije bhije	Rope, cloth wrapped around any object
ukwanezile	To perfect
uloliwe	Train
umamkeliso	Food rations for farmers
umqondiso	A sign, token, evidence, indication
uqothile	Porridge made from the dried sediment of xhosa beer
uswelekile	He/she died
wonile	He/she did wrong

[E/O…i/u…a]

ebalekiswa	He/she is caused to run
ebebheculula	He/she opened the eyelid with the finger
ebephengulula	He/she turn upside down
esekenziswa	He/she is caused to work
goduka	Go home
khohlelisa	Cause to cough
ndiyapendula	I answer
uyatyeleliswa	They are caused to visit
ukwaneliseka	To satisfy
uxleliswa	He/she is caused to tell
uyabalekiswa	He/she is caused to run
uyafeketiswa	He/she is caused to play/trifle

uyakyeleliswa	------------
uyalesiswa	He/she is caused to read
uyapenduliswa	He/she is caused to answer
uyaqondiswa	He/she is caused to understand
uyasekenziswa	They are caused to work
uyathengiswa	He/she is caused to buy
uyathetiswa	He/she is caused to speak
wathengiswa	They are caused to buy
wenziswa	He/she is caused to make
wonziswa	-----------

[E/O…a…i]

ebebhadekile	He was caught off-guard
ebebhadlile	He was caused to become angry
ebebhankroti	Bankrupt
ebeqabelisa	Cause to ascend or go over the top
ebeyalekisa	To teach one lesson
ebhakontini	Outdoor brick made oven
ebhalkonini	Balcony
emdlalweni weentonga	In stick-fighting
etabeni	Mountain
ebeqabukelisa	To look back
esandleni	Applause
impelaveki	Weekend
isiqedaveki	Potent alcoholic drink made from mouldy bread and yeast
Okwe baloni igqabhuka	To burst like a balloon
oomalume	Uncles
ulwandlekazi	Ocean
unesandlozi	Wide starring eyes

[E/O…a]

ebebaleka	He/she runs
enza	Make
feketha	Play/trifle
godola	Feel cold
irholoma	Eye socket
khohlela	Cough
khokela	Lead
kondlaliwe	The bed is ready
leja	That (one) thing
lena	This (one) thing
lona	That (one) person
lowa	This (one) person
ndenza	I make
ndiyabaleka	I run

ndiyabona	I see
ndiyafeketha	I play/trifle
ndiyalesa	I read
ndiyaqonda	I understand
ndiyasebenza	I work
ndiyathenga	I buy
Ndiyathetha	I speak
ndiyeva	I hear
oobawo	Fathers
ubongani	Bongani
ukubhalela	To write
ukuqeda	Finish/accomplish
ukuqoma	Feel pain in the mind or heart
ukurhonoza	Dirty, soil, smudge
ukuromensa	Romance
ukusondezela	Bring nearer for a person or reason
ukusweleka	To die
ukuthoba	To degrade
ukwaneza	To satisfy/make perfect
Ulwahlulelwano	Act of sharing/fellowship
umbabebabe	Skin irritation
uyakyelela	---------------
uyathenga	He/she buys
uyatyelela	He/she visits
uyeva	He/she hears
yenza	He/she makes

4. The theory as applied to isiXhosa

The constraints applied in Pulleyblank's (2002) No disagreement theory are adopted as a crucial set of constraints which have the capacity to drive harmony. As the title suggests harmony is achieved through the prohibition of disagreeing segments (Pulleyblank 2002). The basic theory presents a finite set of constraints, including harmony constraints prohibiting a sequence of opposite features or feature values. As well as markedness constraints which prohibit the occurrence of a specific feature value or particular segment. Finally, faithfulness constraints are included in order to enforce some correspondence between the input and output representations. However, prohibiting an adjacent sequence of opposite segments is not always sufficient to produce the complex harmony patterns attested in different languages. Therefore, other constraints are also necessary. In general each constraint must be selected based on its effectiveness in directly realising the harmony pattern being analysed.

4.1. Constraint definitions

4.1.1. Markedness constraints:

(1) *-ATR,+Hi (Archangeli & Pulleyblank 1994, Bakovic 2000, Kramer 2001)
For each vowel 'V' in the output, assign a violation if: V is [-ATR] ; V is [+hi]

This markedness constraint functions as a restriction on all high vowels for the feature value [-ATR]. Therefore, all high vowels must retain the feature value [+ATR]. The isiXhosa vowel inventory does not include harmonic counter parts for the high vowels [i,u]. The attested candidate is therefore prohibited from containing the high retracted vowels [ɪ, ʊ]. This particular markedness constraint ensures that the high vowel always occurs with the same feature value for ATR.

(2) *+ATR,+Lo (Archangeli & Pulleyblank 1994, Bakovic 2000, Kramer 2001)
For each vowel 'V', assign a violation if: V is [+ATR] ; V is [+lo]

The markedness constraint above limits the occurrence of the low vowel. This constraint states that a vowel cannot be both low as well as [+ATR]. Therefore, the vowel inventory of isiXhosa lacks a harmonic counterpart for the low vowel [a] as well as the high vowels. The advanced low vowel [ə] is hence prohibited as it does not form a part of the vowel inventory. The attested candidate must occur with the retracted low vowel [a].

(3) *+ATR (Archangeli & Pulleyblank 1994, Bakovic 2000, Kramer 2001)
For each vowel 'V', assign a violation if: V is [+ATR]

The markedness constraint above prohibits the advancement of any vowel [i,u,e,o]. Therefore, a candidate containing a vowel with the feature [+ATR] will receive a violation for each vowel with this feature value. The candidate with the most segments retaining this feature value will be considered the worst choice by this constraint.

(4) *-ATR (Archangeli & Pulleyblank 1994, Bakovic 2000, Kramer 2001)
For each vowel, 'V' assign a violation if: V is [+ATR]

Here the prohibited feature value is [-ATR] which states that no vowel may be retracted. Violations are assigned to all segments with the feature value [-ATR]. This includes [ɪ,ʊ,ɛ,ɔ], these vowels are prohibited by this particular markedness constraint.

4.1.2. Harmony constraints:

(5) *-ATR-C-+ATR (Pulleyblank 2002)
For each vowel sequence, 'V-C-V' assign a violation if: V[-ATR] is followed by V[+ATR]

- This sequence may have an intervening consonant between adjacent segments.
- The directionality specified in this constraint states that the [-ATR] segment must precede the [+ATR] segment.

The constraint above prohibits the sequential occurrence of two opposite values for the same feature. This may be described as either a harmony or spread constraint, the basic function however is that it prohibits disagreement between segments for the feature ATR. Any adjacent sequential vowel segment must not have a [-ATR] vowel followed by a [+ATR] vowel. According to this constraint it is more acceptable for a vowel with the feature [-ATR] to be preceded by a vowel with the feature [+ATR]. This particular constraint also places a parameter on distance. As a constraint of medium distance based on the definition by Pulleyblank (2002), this prohibition only applies to vowel segments separated by a single consonant. Therefore, violations are not assigned for the total number of segments bearing opposite values for the same feature, but to the candidate with the higher number of local segments with opposite values. This distance parameter therefore favours harmony cases with opaque segments.

(6) *-ATR∞+ATR (Pulleyblank 2002)
For each vowel sequence, 'V' assign a violation if: V[-ATR] is followed by V[+ATR]

- There are no parameters specified on the distance between these segments.
- There are no parameters specified on the intervening segments between these vowels.
- The directionality specified in this constraint states that the [-ATR] segment must precede the [+ATR] segment.

The constraint above is another which can be described as either a harmony or spread constraint type. With the same basic sequential prohibition this constraint also states that a [-ATR] segment may not precede a segment with the opposite value [+ATR]. The same directionality requirement applies as the constraint specifies which feature value may not precede the other. However, the distance specification is how we distinguish between the functions of these constraints. They do not perform the same role in the evaluation of candidates as this constraint assigns violations to sequential segments whether they are adjacent or a distance apart. Therefore, while the previous version would only assign violations to vowels separated only by a consonant, this constraint assignments violations for all the sequential opposites within a candidate. Hence, cases in which we find transparency are favoured by this constraint.

(7) *+ATR-C--ATR (Pulleyblank 2002)
For each vowel sequence, 'V-C-V' assign a violation if: V[+ATR] is followed by V[-ATR]

- This sequence may have an intervening consonant between adjacent segments.
- The directionality specified in this constraint states that the [+ATR] segment must precede the [-ATR] segment.

Also a harmony/spread constraint, with the same basic function as the two discussed above, this constraint prohibits a sequence of segments which disagree for the feature ATR. A violation is assigned when two vowel segments separated by a consonant share opposite values for the same feature. In this case the directionality is still specified but requires the opposite sequence to the initial two harmony constraints. This constraint therefore prefers cases in which a [+ATR] segment is preceded by a [-ATR] segment. This harmony constraint also places a parameter on distance by only assigning violations to those vowel segments with opposite features, which are adjacent except for an intervening consonant. Hence, this is also described as a local harmony constraint, which favours cases of opacity to transparency.

(8) *+ATR∞-ATR (Pulleyblank 2002)
For each vowel sequence, 'V' assign a violation if: V[+ATR] is followed by V[-ATR]

- There are no parameters specified on the distance between these segments.
- There are no parameters specified on the intervening segments between these vowels.
- The directionality specified in this constraint states that the [+ATR] segment must precede the [-ATR] segment.

The final harmony constraint included in this ranking performs the same basic function as the one described above. Sequences of [+ATR] followed by [-ATR] are prohibited. Therefore, it reflects a sequential prohibition, with a directionality component implied in the order in which the featured segments are prohibited from occurring. As previously discussed there is a local and a distance version of each harmony constraint. This particular version assigns violations to sequences of opposite segments across the entire word, and not just adjacent segments. It therefore favours cases of transparency in which the majority of the segments will share the same feature value even though one might not.

4.1.3. Two feature, Local Harmony constraints:

The two-feature harmony constraints to follow have not been adopted from Pulleyblank's (2002) sample constraint set as the constraints above. The following harmony constraints have been adapted based on the concept of correspondence (Rose & Walker 2004, Walker 2009), and similar constraint proposals designed to consider more than one feature simultaneously (Walker 2016).

(9) *-ATR/-LO-C-+ATR-LO
For each vowel sequence, 'V-C-V' assign a violation if: V[-ATR,-LO] is followed by V[+ATR,-LO]

- This sequence may have an intervening consonant between adjacent segments.
- The directionality specified in this constraint states that the [-ATR,-Lo] segment must precede the [+ATR,-Lo] segment.

This harmony constraint and the following are particularly significant and essential to the ranking of the isiXhosa harmony pattern. In the same way as those before, this constraint places a sequential prohibition on specific combinations of features. The targeted sequences in this case share not one but two specified features. Therefore, the preceding segment must be specified as both [–ATR] and [–Low] in order to activate this constraint, it is also essential that the following segment be specified for both features [–ATR] and [–Low]. The sequential prohibition now requires an interaction between two binary features. The constraint will therefore only be activated if both segments share the feature [–Low], with the preceding segment specified for [-ATR] and the following segment for [+ATR]. This type of constraint is different from the original no-disagreement constraint, as it also checks for agreement.

(10) *+ATR/-LO-C-ATR-LO

For each vowel sequence, ‘V-C-V’ assign a violation if:

a. V[+ATR,-LO] is followed by V[-ATR,-LO]

b. This sequence may have an intervening consonant between adjacent segments.

c. The directionality specified in this constraint states that the [+ATR,-Lo] segment must precede the [-ATR,-Lo] segment.

This is the second constraint which specifies requirements for two binary features. As a harmony constraint the prohibition here is placed on a specific sequence of features. The sequence of segments must begin with a vowel specified for [+ATR,-Low], followed by a vowel specified for [-ATR,-Low] in order to activate the constraint. As the one above this constraint includes specifications for two features, unlike the initial harmony constraints which include a specification for only the ATR value of the preceding and following segment. As this constraint is the exact reverse of the previous harmony constraint, the directionality requirement is different. This constraint will only assign a violation if the [+ATR,-Lo] segment is the preceding segment.

4.1.4. Faithfulness constraint:

(11) Ident[ATR] (McCarthy & Prince 1995)

For each vowel segment, ‘V’ assign a violation if:

a. Input V is [-ATR], and output V is [+ATR]

b. Input V is [+ATR], and output V is [-ATR]

c. The input and output vowel segments do not share the same ATR value.

This constraint is the only faithfulness constraint included in the ranking. It requires a relationship between the input and output segments specifically their value for the feature ATR. Therefore, whatever value for ATR a segment has in the input form, this value must be retained in the output form.

4.2. Optimality Theory component GEN

Optimality theory serves as a method of formalisation, through which the exigencies specific to a particular language are represented as language internal markedness generalisations (Prince and Smolensky 2004). These generalisations are used to inform the construction and selection of constraints, which are then ranked to ensure the linguistic output will reflect the grammar of a specific language. According to literature discussing the emergence of OT there are three distinct parts to the process, these include GEN, CON, & EVAL (Smolensky 2005).

4.2.1. Features

When analysing vowel harmony we are analysing the pattern in which a feature is shared across multiple vowel segments within a specific domain. This feature is generally a phonological property, in the case of isiXhosa it is either the advancement or retraction of the tongue root. The harmony pattern exhibits the feature advanced tongue root (ATR) occurring only as a result of vowel raising, while the feature retracted tongue root (RTR) appears in all other situations (Jokweni & Thipa 1996). The marked constraint is therefore used to represent both features in a binary constraint [+-ATR]. The feature [+-ATR] is used to describe the phonetic distinction between the mid-vowels that appear in different phonological contexts in isiXhosa. The relevant features to isiXhosa vowel harmony are therefore considered to be advanced tongue root as well as height. According to previous findings ATR is identified as the shared feature across the vowel segments within a word. Considering the influence of vowel raising as well as the limitations on which features a specific vowel may retain, there is also an interaction with the feature height. Therefore, the general pattern is characterised as one of ATR harmony, however in order to structurally define this pattern one must incorporate its relation to the feature height. If the feature [+ATR] is only activated when a high vowel trigger appears to the right of any preceding mid-vowels, but the same effect is not created by a mid-vowel with a [+ATR] value, then the height of the triggering segment becomes essential. As just described all of the vowels perform a specific function in isiXhosa, the mid-vowels are the undergoers, as they consistently alternate depending on the phonological context. The high vowels serves as triggers for the feature [+ATR]. The low vowel however is opaque, it has no harmonic counterpart with the value [+ATR] and acts as harmony blocker requiring that every non-high vowel preceding it remain [-ATR].

4.2.2. IsiXhosa vowel matrix:

(12) Vowel feature matrix:

	i	e	ɛ	a	o	ɔ	u
high	+	-	-	-	-	-	+
low	-	-	-	+	-	-	-
ATR	+	+	-	-	+	-	+

Presented in the tableau above are the vowels which appear in the isiXhosa inventory. According to previous literature (Jordan 1966, Pahl 1967) as well as my own acoustic findings, there are seven vowels. These include the basic five high front /i/, high back /u/, mid-front /e/, mid-back /o/ and the low vowel /a/. The two additional vowels include the [-ATR] front mid-vowel /ɛ/ and the [-ATR] back mid-vowel /ɔ/. Based on the inventory the high and low vowels do not have the same harmonic counterparts for the feature ATR, and are therefore unable to alternate. However, because this feature is not contrastive the incomplete harmonicity does not result in any limitation to the possible meanings of a lexical item. The interest in this pattern is therefore purely related to the way the language functions internally, and how the articulation and production of sounds are made easier. The alternation for the feature ATR is therefore viewed as a method of minimizing the instances of inertia.

4.2.3. Selection of Input form

The general concept of the input form in phonology can be defined as the underlying phonetic structure of the word before any morphological or phonological change occurs. However, this underlying structure is not always easily determined and may be realised differently according to various phonologists, due to the manner in which an alternation is interpreted or analysed. Hence, in order to contribute to a more comprehensive and uniform OT approach it is best to entirely eliminate the necessity of the input form. This task may be accomplished by terminating the reliance on any constraints which place violations on the relation between possible candidates and the input form, however it may also be accomplished in the GEN process (Smolensky 1996). By including all possible realisations of each vowel for the feature [ATR] in a number of input forms, this eliminates the relevance of the input form in the analysis process entirely. If all candidates are considered as the underlying form in the analysis, then no form is technically adopted as the input. Therefore, although a default feature value is established in chapter six discussing the acoustic analysis, this form will not be presented as the predetermined input in the OT analysis. This decision is not to discredit the phonetic results of the alternation elicited from the acoustic analysis, but rather to initiate an OT approach which may be considered more uniform in that the resultant optimal candidate will not be dependent on a predetermined and idiosyncratic underlying form. The number of input forms should correspond with the number of candidates, this will be determined with the simple formula; number of candidates = 2^N. The number of candidates will therefore be dependent on the number of vowels in the lexical representation, each realised with both values for the feature [ATR]. The formula 2 to the power of N therefore reflects 2 multiplied by itself according to the number of vowels in the lexical representation.

(13) Example: komzi 'feeling of community'

Input: komzi Cset: 1. komzi 2. komzɪ 3. kɔmzɪ 4. kɔmzi	Input: komzɪ Cset: 1. komzi 2. komzɪ 3. kɔmzɪ 4. kɔmzi	Input: kɔmzɪ Cset: 1. komzi 2. komzɪ 3. kɔmzɪ 4. kɔmzi	Input: kɔmzi Cset: 1. komzi 2. komzɪ 3. kɔmzɪ 4. kɔmzi

4.2.4. Richness of the base

The universality of GEN is reflected in the candidate forms emitted, which should include those attested in all languages. This universal property applies to the candidates as well as the input forms, and is referred to as the freedom of analysis or inclusivity of GEN (McCarthy 2002). Although this theory should be limited, excluding the primitive structures that are essential to the language. Hence, in the case of isiXhosa this would imply that it is not necessary to include candidates with segments which are [+ATR,+Lo] or [-ATR,+Hi] as such vowels do not exist in the inventory. However, as there are languages containing such vowels in their inventory, the examples above would have to be included in order for the approach to be aligned with the richness of the base hypothesis. This would involve the perception that the input which is generally considered the underlying form, is actually universal and the set of potential inputs should be the same for all languages (McCarthy 2002). Therefore, the vowels which do not exist in the vowel inventory of isiXhosa must still form part of the input set, not to suggest that such forms exist in the language but rather to contend with other inputs in order to accurately define the isiXhosa speaker's knowledge of their language. To place any restriction on the input form would result in a less effective interaction between the markedness and faithfulness constraints, as the operation of the ranking would depend on specific underlying postulations. The elimination of input forms not applicable to the language should be left to constraint interactions and ranking (Smolensky 1993). The overall aim is to align the approach taken in this paper with the richness of the base hypothesis, to escape any reliance on an input form and avoid some of the conceptual and empirical problems faced in previous research (Kurisu 2000).

4.2.5. Candidate selection

The general harmony pattern in isiXhosa exhibits an alternation in all mid-vowels preceding a high vowel (Jokweni & Thipa 1996). Hence, the only instances in which the vowel raising phenomenon would occur is with the attachment of suffixes. If a high vowel is contained in the suffix it will certainly affect the [ATR] value of the preceding mid-vowels, where as a prefix containing a high vowel has no effect on the mid-vowels that follow. Therefore, the pattern also exhibits regressive spreading, all of which must be considered in the selection of data. Hence, the harmony process that occurs in isiXhosa is quite complex requiring that the system of analysis account for a number of exigencies. As in the Pulaar harmony system (Archangeli & Pulleyblank 2002, Sasa 2009) the OT approach used to analyse isiXhosa must be capable of dealing with the directionality of the pattern, as well as the sour grapes problem in the cases in which only partial harmony can be achieved (McCarthy 2011, Sasa 2009). All attested cases of harmony in isiXhosa must be analysed according to the same constraint ranking in order to determine the functional motivation of each constraint, as well as to ensure the OT approach adopted has the capacity to deal with all unattested candidates. The candidates included in GEN should reflect attested examples of all possible harmony cases in isiXhosa, this includes total harmony, partial harmony, directionality as well as cases of opacity/blocking. The harmony cases above may be analysed within the following morphological categories.

(14) **Total Harmony:**

ndɔnɛ ‘I did wrong’ /ona/

komzi ‘feeling of community’

(15) **Partial harmony:**

ebonisa ‘he/she is caused to see’ /bona/

goduka ‘go home’ /goduk/

(16) **Directionality:**

esilɛ ‘donkey’

oothixɔ ‘gods’

(17) **Blocking:**

ɛtabeni ‘mountain’

ɔmalumɛ ‘uncle’

The examples above reflect the surface representation of each word and therefore reflect the phonetic effects of any morphological transformation. The cases of total harmony include lexical representations in which all segments share either the value [-ATR] or [+ATR]. Hence, total harmony may be achieved either through the process of vowel raising, or in the absence of any high vowel trigger. Cases of partial harmony include phonological situations in which raising has occurred due to a high vowel trigger, which is then followed by the low vowel [a] causing the first two adjacent segments to agree for the feature value [+ATR] followed by a final vowel which is [-ATR]. The cases exhibiting directionality include a high vowel both preceded by and following a mid-vowel. In such cases only partial harmony should be achieved and directionality must be enforced in the constraint ranking. The final harmony case provided above is blocking, in which the high vowel and preceding mid-vowel are separated by the low vowel [a] in which case harmony cannot be triggered. In order to develop a constraint ranking capable of accounting for the complexities of this harmony system, all possible and attested cases of harmony must be processed. GEN must also however account for what is unattested and not possible in this particular language. Hence, the candidate set must consist of all possible alternations for ATR, regardless of the inventory limitations of the vowel system. Here we refer to the lack of harmonic completeness in the isiXhosa vowel inventory (Smolensky 2005). By accounting for all possibilities, the resulting grammar is one which may derive the harmony pattern of this particular language from any random candidate set or input specification.

4.3. Underlying status for [+ATR] mid-variants

The general hypothesis proposed by Jokweni & Thipa (1996) states that there is a marked mid-variant of the feature ATR in isiXhosa. After establishing a tongue root distinction between two sets of mid-vowels in isiXhosa, it was also determined that one set occur in limited instances, while the other occurs more freely. The [+ATR] mid-variant {e, o} is considered the marked option and is the result of the active spreading of the feature [+ATR]. However, this assimilatory effect cannot be activated by any [+ATR] vowel, the trigger must share both features [+ATR, +Hi]. Therefore, the general hypothesis states that all occurrences of [+ATR] mid-vowels must be attributed to the influence of a high vowel trigger. This concept of high vowel raising, which states that a [+ATR] mid-vowel cannot occur without a following high vowel trigger, is then weakened by data cited in Pahl (1967). This data includes cases of semi-closed or [+ATR] mid-vowels occurring without a high vowel trigger and therefore not as a result of vowel raising. The data presented below include only examples of interjections and ideophones, which it is important to note are likely to vary sociolinguistically.

(18) Semi-closed vowels, No vowel raising (Pahl 1967)

<u>Interjections:</u>

a. yeha ‘expressing shock’
b. yoo ‘expressing surprise’

<u>Ideophones:</u>

c. geqe ‘implying complete independence’
d. nkqo ‘implying uprightness’
e. gqolo ‘implying continuity’

4.3.1. Interjections & Ideophones

Based on this evidence Jokweni & Thipa (1996) conclude that the [+ATR] mid-variant must be inherent to the vowel inventory of isiXhosa. Hence, if it is possible for these vowels to occur without the influence of any other vowel, they must form part of the underlying inventory. According to the literature this data is fully acknowledged, however the effects of high vowel raising remain the prominent cause for the occurrence of the [+ATR] mid-variant.

It is important to note however as mentioned briefly above, that lexical items such as ideophones and interjections are highly dependent on social factors. The lexical items used for these emotive and expressive purposes will differ not only cross-linguistically but also between different dialects and linguistic communities. These lexical items exhibit a strong connection between sound and meaning in all languages, and therefore cannot be interpreted as a typical example of the phonetic exigencies. In English the concept of smallness is expressed in high front vowels, while largeness is expressed in the high back vowel space. The same pattern is observed in Russian, with high front vowels expressing smallness and low back vowels largeness (Maduka 1988). Such lexical representations usually represent instances of anomalous phonology, e.g. clicks used for sound-symbolism in English; and unattested consonant clusters (Sapir 1929). Any assumptions made about such data must be considered from a phono-semantic perspective and not as purely phonetic phenomena.

Considering the socially dependent nature of lexical items such as interjections and ideophones it is not surprising that these specific constructions are not found in my own collected data, as my speaker does not share the same dialect as is studied in Jokweni & Thipa (1996) or Pahl (1967). Only one example of an ideophone is found in this data which is, 'Hememe' /hɛmɛmɛ/ meaning "you're in trouble". This single example cannot represent an entire class of lexical items however according to an acoustic analysis of this word and others with only mid-vowels, the feature [+ATR] will not spread without a high vowel trigger. Based on the results of the acoustic analysis the mid-vowels in /hɛmɛmɛ/ all occur as the retracted form. In this particular dialect the acoustic data also reveals that every case of mid-vowel occurrence without high vowel influence appears as the retracted form. It is therefore proposed here that the cases presented in Jokweni & Thipa (1996) reflect anomalous instances. Analysing the default state of mid-vowels exhibits a consistent reversion to the retracted form without the influence of vowel raising.

5. OT Analysis

5.1. Summary of the system

The following constraints include a combination of markedness and faithfulness constraints proposed in Pulleyblank's (2002) no-disagreement approach. The basic local and distance harmony constraints below also form part of the no-disagreement sample constraint set (Pulleyblank 2002), the two-feature harmony constraints however are adapted based on the concept of correspondence (Rose & Walker 2004, Walker 2009), and similar constraint proposals simultaneously considering more than one feature (Walker 2016). A detailed and referenced explanation of the constraints below is provided in section 4.1 above.

5.1.1. Constraint definitions

Markedness constraints:

*+ATR: Assign a violation for every vowel with the feature [+ATR]

*-ATR: Assign a violation for every vowel with the feature [–ATR]

*+ATR/+Lo: Assign a violation for every vowel with both features [+ATR] & [+Lo]

*-ATR/+Hi: Assign a violation for every vowel with both features [-ATR] & [+Hi]

Harmony constraints:

*+ATR-C--ATR: Assign a violation for every vowel sequence with a [+ATR] vowel preceding a [-ATR] vowel, separated only by a consonant between vowels.

*-ATR-C-+ATR: Assign a violation for every vowel sequence with a [-ATR] vowel preceding a [+ATR] vowel, separated only by a consonant between vowels.

*+ATR∞-ATR: Assign a violation for every vowel sequence with a [+ATR] vowel preceding a [-ATR] vowel.

*-ATR∞+ATR: Assign a violation for every vowel sequence with a [-ATR] vowel preceding a [+ATR] vowel.

Faithfulness constraint:

Ident[ATR]: Assign a violation for every vowel in the output with a different ATR value as the corresponding vowel in the input.

Additional two-feature harmony constraints:

*+ATR/-LO-C-ATR-LO: Assign a violation for every vowel sequence with an initial vowel carrying the features [+ATR] & [-Lo] preceding a vowel with the features [-ATR] & [-Lo], sequences separated only by a consonant between vowels.

*-ATR/-LO-C-+ATR-LO: Assign a violation for every vowel sequence with an initial vowel carrying the features [-ATR] & [-Lo] preceding a vowel with the features [+ATR] & [-Lo], sequences separated only by a consonant between vowels.

5.1.2. Constraint functions

The markedness constraints which limit the occurrence of a particular ATR value such as [+ATR]/ [-ATR], these constraints determine which ATR value will serve as the default option. If *[+ATR] is ranked above *[-ATR] then the default form will have the value [-ATR]. The ranking of *[+ATR] is quite significant in generating the correct output form for isiXhosa. It is imperative that this markedness constraint is ranked above *[-ATR] as well as all of the harmony constraints. The second set of markedness constraints relate to two features, taking into account both the ATR value of a segment as well as its value for height. The markedness constraints limiting both features ATR and height are required to eliminate all candidates which do not exist in the vowel inventory of isiXhosa. Vowels which have the feature value [+ATR] must also share the feature [+Hi], there is a strict fundamental relation between height and the feature ATR, see chapter four figure (12). Hence, if a vowel holds the feature value [-ATR] it cannot also share the value [+Hi]. These markedness constraints are extremely important and must be ranked among the highest constraints. All candidates with vowels which do not exist in the vowel inventory should be eliminated based on their violation of the inventory limitations, and not by the harmony constraints.

The harmony constraints focusing only on the feature ATR may be divided into two sets distinguished by distance. The first set of harmony constraints place limitations on vowel sequences with opposite values for the feature ATR at a close distance. Therefore, the only segments which may be counted are those sequences in which the two vowel segments are separated by a consonant and not another vowel. E.g. the [+ATR] segment must be directly adjacent to the [-ATR] segment in order for the violation to be counted. These local distance harmony constraints ensure that cases of opacity will be accounted for. If the focus is always on the total number of segments with shared features and not the number of adjacent segments with shared features, then the system would not have the capacity to select the attested candidate in cases of blocking. The second set of harmony constraints are not local but apply to segments even at a distance. These constraints therefore limit the occurrence of no-disagreement for the feature ATR across an entire word. The same sequences are prohibited as in the local harmony constraint, however the vowel sequences need not be directly adjacent in order to violate the constraint. By counting all instances of opposing values for ATR, a case of total harmony is preferred above partial harmony. However, this does not mean the attested candidate will reflect a case of total harmony, it simply means transparency may occur as the total number of segments sharing the same feature value within the word is valued above the number of adjacent segments sharing the same feature value. With these two constraints the system can accommodate the exigencies of both opacity and transparency in the harmony pattern.

The additional local harmony constraints are of particular significance. These are the only harmony constraints which consider more than one feature at a time. Hence, these constraints will not be activated by a vowel segment unless it retains both features specified in the constraint. The adapted harmony constraint considers vowel sequences with segments sharing a specified value for ATR as well as a height requirement. As we've already included four distinct harmony constraints limiting the occurrence of vowel sequences with segments which disagree for ATR, the feature of particular interest here is height. The addition of a height requirement on these sequential prohibitions adds the final requisite for the generalised harmony pattern of isiXhosa. Within the generalisations for producing harmony, the trigger which causes vowel raising has very distinct featural requirements. The trigger must share the feature value [+ATR] as well as [+Hi]. Vowel raising will therefore never occur as a result of a [+ATR/-Hi] vowel. Without accounting for this predisposition, the optimal candidate would be one in which raising is accounted for by the general sequential prohibitions for ATR.

The final and lowest ranked constraint is the faithfulness constraint. This constraint requires that the ATR value of a vowel in the input form be identical to that of the corresponding vowel in the output form. A connection is formed between the vowel segments in the input and those in the output. This relationship requires symmetry between each input vowel and its output counterpart in terms of their value for ATR. The faithfulness constraint is the lowest ranked constraint because all markedness and harmony constraints must be ranked above it in order to allow phonological alternations to take place. Any alternation such as vowel raising would be hindered by the faithfulness restrictions. Therefore, to accommodate the harmony system, the faithfulness constraint must be outranked by all other constraints prohibiting disharmony and thereby driving the correct harmony pattern.

5.2. Input form & Candidate selection

5.2.1. The feature

The feature of relevance in the vowel harmony system of isiXhosa is tongue root position, specifically the binary feature ATR. This feature is only relevant to mid-vowels, as they are the only vowels in the inventory which have harmonic counterparts for the feature ATR. Based on the acoustic analysis the feature value [+ATR] is the active feature, spreading onto mid-vowels only when the appropriate trigger is present. The OT constraint ranking must account for spread of [+ATR] while respecting the occurrence of [-ATR] as the default option. However, because the necessary trigger holds both features [+ATR] & [+Hi] the constraints must consider the feature height when determining where raising should or should not occur.

5.2.2. The vowels

The alternating segments within this particular harmony pattern are limited to mid-vowels. Mid-vowels serve as the undergoers, when a sequential prohibition constraint requires the alternation of a segment to either the [+ATR]/ [-ATR] form, that alternating segment will always be a mid-vowel. The reason for such sequential prohibitions is to accommodate the process of vowel raising in isiXhosa. The constraints restricting sequential disagreement for ATR are the driving force for ATR harmony. If the harmony constraint *-ATR∞+ATR is ranked above *+ATR∞-ATR, then the optimal candidate will be one in which we see vowel raising and [+ATR] harmony between segments. However, such a candidate may not correspond with the attested form in isiXhosa as raising is not triggered by any vowel with the feature [+ATR]. Therefore, the requirements on the triggering segment include more than a specific value for the ATR. The function of high vowels as the trigger for vowel raising are therefore represented by the harmony constraints which consider both the feature ATR as well as height. The final vowel is the low vowel [a] which has neither an ATR counterpart nor a low front or back counterpart. The low vowel blocks the iteration of vowel raising. Therefore, depending on its position within the lexical item, the intervening low vowel will block the assimilation of all mid-vowels preceding it.

5.2.3. The input form

If the purpose of OT is to derive the grammar of one language from the universal characteristics of many, then regardless of what the underlying form may be, the system itself must not depend on a particular input to successfully derive the correct pattern. Therefore, the inputs included will be a reflection of every possible candidate. The number of inputs must correspond with the number of candidates these are therefore constructed using the formula 2^N. This formula ensures the calculation of two realisations for each vowel in a specific lexical item. By using this formula the input form will include all possible realisations of this alternation for ATR.

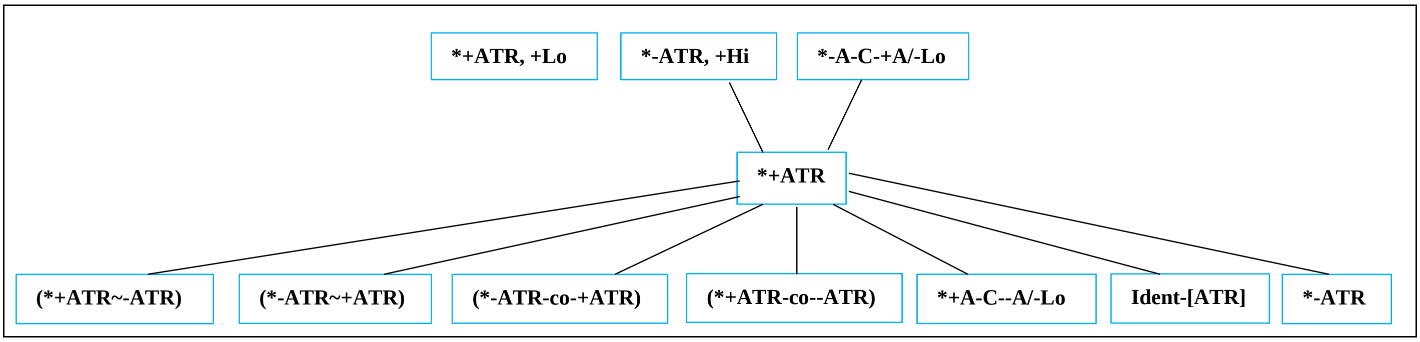
5.2.4. Candidate selection

The candidate selection should not only include all possible realisations of each vowel for the feature ATR but must also include all attested forms found in isiXhosa. As we derive the input forms to include in the analysis, the selection of candidates must be calculated in the same way. Every possible realisation of each vowel for the feature ATR must be represented in the candidate set. We have therefore created a system in which every candidate has competed against all other possible candidates through the inclusion of all possible forms in both the input as well as the candidate set.

5.3. Constraint ranking

The constraints outlined above in chapter four form the initial part of the CON aspect of the OT approach. Each constraint listed in the outline performs an essential function in driving harmony according to the exigencies of this language. In order for these constraints to perform effectively they need to be prioritised in a particular way, this makes up the EVAL aspect of OT. The ranking of these constraints is imperative to the function of each constraint. E.g. If the harmony constraint prohibiting *-ATR-C-+ATR is ranked below the faithfulness constraint Ident[ATR], then the constraint restricting sequential prohibition will not be allowed to force any alternation in order to drive harmony. The optimal ranking will allow each constraint to perform the function for which it has been included. The necessary ranking for deriving the optimal candidate in isiXhosa is dependent on the type of candidate and resultant harmony. The entire ranking proposed for driving this harmony pattern is only necessary in dealing with specifically complex candidates. In certain cases only a portion of this ranking will be necessary to derive harmony.

(1) Overall ranking:



The ranking above represents the entire filtering system needed to derive the harmony pattern for isiXhosa. The highest ranked constraints include the markedness constraints eliminating all candidates containing vowels which do not exist in the inventory. The highest ranked harmony constraint is the sequential prohibition on the occurrence of a [-ATR] segment preceding a [+ATR] segment, with the addition of a height restriction. This constraint ensures that the only vowels capable of causing raising or regressive assimilation for [+ATR] are vowels which share both features [+ATR] & [+Hi]. The critically ranked constraint is the markedness constraint *[+ATR], this is the only constraint that must be ranked above and below other constraints. This ranking does not require that the highest ranked constraints described above be ranked in a particular order amongst each other. The lower ranked constraints may also be ranked in any order amongst each other. The only constraint both above and below particular constraints is the constraint prohibiting the occurrence of the feature value [+ATR]. This constraint performs a vital function in ensuring that the vowels unaffected by a trigger, occur as [-ATR] in the output form. Although the harmony pattern itself is a complex one derived from a set of generalisations presenting the exigencies of how the feature ATR is shared between vowels, the constraint ranking needed to derive this type of harmony is quite transparent and straightforward.

5.4. Simple example

Particular lexical items provide easy cases for deriving harmony. If a word consists only of mid-vowels, all that needs to be done is for the default option to be activated. The markedness constraints limiting the occurrence of particular values for ATR are vital in this situation, because if two candidates proceed through the ranking and equally fulfill the harmony criteria, the deciding factor will be the kind of harmony required in that particular case. If the word consists only of mid-vowels then the optimal candidate should have total [-ATR] harmony.

(2) Basic harmony interaction:

/ndenze/ ‘I made’

/ndenze/	*+ATR	*+ATR∞-ATR	*-ATR∞+ATR	*-ATR
a. ndenze	*			
b. ndɛnze			*	
c. ndɛnzɛ ☺				*
d. ndenzɛ		*		

The optimal candidate in the tableau above represents a case of total harmony. The lexical item has a total of two vowels which means they each have two forms for the feature ATR. This equates to a total of four candidates representing every possible variation of ATR within the entire word. Two candidates represent cases of partial harmony, the other two are cases of total harmony. The attested candidate in this case with a lexical item consisting only of mid-vowels, should be a case of total harmony. This means that candidates (2b) & (2d) could not be the attested form and need to be eliminated. Candidate (2b) is eliminated by the harmony constraint restricting the occurrence of a [–ATR] segment before a [+ATR] segment. Candidate (2d) is eliminated with the opposite harmony constraint restricting the occurrence of a [+ATR] segment before a [–ATR] segment. The rapid consecutive elimination of these sub-optimal candidates exhibits an unnecessary hierarchical relation between these constraints. Therefore, ranking the second harmony constraint above the first will not result in a different outcome, both candidates will still be the first to be eliminated by the ranking. The deciding constraints here are the markedness constraints. Candidate (2a) & (2c), both reflect cases of total harmony, the decision now becomes which harmony is preferred. To derive the harmony pattern in isiXhosa, the markedness constraint *+ATR must be ranked above the markedness constraint *-ATR as seen above. Thereby candidate (2a) which reflects total [+ATR] harmony is eliminated, leaving the optimal candidate and attested form, which is a case of total [–ATR] harmony. In this particular situation only the distance harmony constraints are needed to eliminate cases of partial harmony. While the markedness constraints restricting the occurrence of certain values for ATR are needed to elicit the default value, in cases where raising does not apply.

5.5. How ranking deals with different harmony cases

As discussed above, the harmony system in isiXhosa is quite complex with numerous requirements. The resultant harmony cases are therefore not consistent and will include instances of total harmony, partial harmony, and even no harmony. The kind of harmony derived is dependent on the make-up of the lexical item in question. Therefore, the example construction analysed above resulted in a case of total harmony, because it only contains mid-vowel segments, there are no other segments influencing their ATR value. In any instance in which there is no high vowel trigger in a word, all of the mid-vowels will revert to the default [–ATR] form which will result in a case of total harmony. However, if a high vowel does occur in the word, even if it does not act as a trigger it will result in a case of partial harmony because a vowel cannot share the features [+Hi] & [-ATR]. The simplest harmony cases are therefore when the lexical item consists of only mid-vowel segments, such as the example above /ndɛnzɛ/. Other cases of total harmony include instances in which the final segment of a lexical item is a high vowel and all preceding segments are either mid or high vowel segments. In these instances all mid-vowels will assimilate for the feature [+ATR] creating total harmony with any high vowels as they already share the feature [+ATR].

More complex cases include lexical items in which a high vowel acts as a trigger but cannot raise all mid-vowels due to its position, and the restriction on directionality. In this case the resultant harmony will be an instance of partial harmony in which the mid-vowels which have not been influenced by the trigger will remain in the [-ATR] form. Partial harmony is also derived when the spreading of [+ATR] by the high vowel trigger is blocked by an intervening low vowel. In this situation the high vowel may affect any preceding mid-vowels, however where the low vowel intervenes the spreading rule is inhibited and all preceding mid-vowels will revert to the [-ATR] form. In any lexical item in which both high and low vowels occur, the resulting harmony will always be partial. This is because high and low vowels share opposing values for the feature ATR and are unable to alternate, as they lack the harmonic counterparts shared by the mid-vowels in the inventory. These instances of partial harmony may be cases in which adjacent segments share a particular value for ATR until a certain point, with the remaining vowels sharing the opposite value. This would be a case of opacity in which the intervening vowel has an iterative effect on any preceding or following vowels. However, there are also instances of partial harmony in which the intervening vowel with a conflicting ATR value does not affect any other vowels and thereafter allows harmony to continue. These cases reflect instances of transparency, in which the conflicting vowel intervenes but is skipped over allowing harmony to continue.

Finally the cases of no harmony are limited, these are lexical items which include a variety of high and low vowels alternating between mid-vowels throughout the word. In these cases harmony is simply not possible due to the position of each vowel. Usually in lexical items like this the high vowel will be directly preceded by a low vowel in which case [+ATR] cannot spread onto any preceding mid-vowels. It will also be the case that the high vowel might not be in the word final position and either the low vowel or mid-vowel following it will share the opposite ATR value. The harmony case will therefore be a reflection of the constituents of the lexical item analysed. Hence, specific constructions will be used to explain each of the harmony cases discussed above.

5.5.1. Total harmony

Instances of total harmony may only be derived from the analysis of particular lexical items. As discussed above, these include words containing only mid-vowel segments, or words which have a high vowel trigger in the word final position preceded only by mid-vowels or high vowels. We will examine these cases of total harmony in the form of constraint tableaux below, using one example of each type of lexical item. The word ‘ndone’ ‘I did wrong’ will be used to display total harmony derived from a lexical item consisting only of mid-vowel segments. The word ‘komzi’ ‘feeling of community’ will be used to exhibit a case of total harmony derived from lexical constructions with a high vowel trigger in the word final position preceded by mid-vowel segments.

(3) Total Harmony: Construction with only mid-vowel segments

/ndone/ ‘I did wrong’

/ndone/	*+ATR	*+ATR∞-ATR	*-ATR∞+ATR	*-ATR
a. ndone	*			
b. ndɔne			*	
c. ndɔnɛ ☺				*
d. ndonɛ		*		

The tableau above displays a case of total harmony derived from the lexical item /ndɔnɛ/. As in the first example the number of candidates is calculated according to the number of vowels. Only four candidates are needed to cover all the possible realisations of ATR for each vowel. Two of these candidates represent cases of partial harmony, in which one of the mid-vowels possess the feature value [-ATR] while the other is [+ATR]. These candidates (3b) & (3d) are consecutively eliminated by the sequential prohibitions of the distance harmony constraints. The candidates reflecting cases of total harmony are then filtered by the markedness constraints. Candidate (3a) being an instance of total [+ATR] harmony, and candidate (3c) [-ATR] harmony. The harmony constraints therefore pushed for total harmony instead of partial harmony eliminating candidates with any sequential disagreements for ATR. The choice between the two cases of total harmony is determined by the markedness constraints prohibiting the occurrence of a particular value for ATR. As the default form for ATR is [-ATR] in isiXhosa, candidate (3a) is eliminated by the highest ranked markedness constraint. This leaves candidate (3c) as the optimal candidate which is also the attested form. The ranking of the markedness constraints are vital in deriving the correct case of harmony. If the constraint *[-ATR] were ranked above *[+ATR], candidate one would be the optimal form which would be incorrect. The ranking must drive the correct choice in situations requiring the default form. The constraint ranking required for this type of construction is simple, harmony constraints are needed to drive ATR agreement between the mid-vowel segments ,and markedness constraints are needed to ensure the resultant harmony is agreement for [-ATR].

(4) Total Harmony: a word-final high vowel, preceded only by a mid-vowel

/komzi/ ‘feeling of community’

/komzi/	*-ATR /+Hi	*-ATR-C-+ATR/-Lo	*+ATR	*+ATR∞-ATR	*-ATR∞+ATR	*-ATR
a. komzi ☺			*			
b. kɔmzi		*			*	*
c. kɔmzɪ	*					*
d. komzɪ	*			*		*

Displayed above is a case of total harmony derived from the lexical item /komzi/. Here harmony is not a result of mid-vowels reverting to the default form, but of mid-vowel raising.

Total harmony is possible because the high vowel trigger is in the word final position allowing it to spread [+ATR] onto any preceding mid-vowels without any inhibiting segments. Unlike the word /ndɔnɛ/ analysed above, /komzi/ includes both a mid-vowel as well as a high vowel. As there are still only two vowels in the word the number of possible realisations of ATR can be represented by four candidates. However, these candidates now include a [-ATR] form of the high vowel, a form which does not exist in the vowel inventory. Therefore, again we have a situation in which there are two candidates representing cases of total harmony, one of which must be eliminated as it includes a [-ATR/+Hi] segment. The first candidates to be eliminated should be those with only partial harmony candidates (4b) & (4d). However, the first candidates eliminated are (4c) & (4d), both violating the highest ranked constraint *-ATR/+Hi. This constraint is specifically used to ensure that harmony cannot be derived with a segment which does not exist in the inventory such as a vowel with the features [-ATR/+Hi]. In the process however we have two kinds of candidates eliminated, one with partial harmony and the other with total harmony both including the segment prohibited by the first constraint.

Left with candidates (4a) & (4b), an additional constraint is now required to force vowel raising in the presence of a high vowel trigger. Candidate (4b) reflects a case of partial harmony, in which the preceding mid-vowel has not assimilated for [+ATR]. This sub-optimal candidate must be eliminated before the markedness constraint *[+ATR] which is critically position higher than the basic distance harmony constraints. In order for candidate (4a), the attested form to win, a constraint must be inserted above the markedness constraint. This constraint must extend further than the basic harmony constraints, by ensuring that harmony is always triggered by the occurrence of a high vowel. Hence, the insertion of the harmony constraint *-ATR-C-+ATR/-Lo, states that it is more important that all mid-vowels assimilate for [+ATR] before a high vowel, than it is to have the optimal candidate have the least number of [+ATR] segments.

In tableau (4) two additional constraints are needed to derive the correct harmony pattern. The candidates with partial harmony were eliminated by the harmony constraints in both cases, however in tableau (4) one partial harmony candidate was eliminated by the first constraint restricting a particular segment from the inventory. This candidate would still have been eliminated by one of the harmony constraints as the ranking pushes for total harmony, however with the addition of the markedness constraint restricting [-ATR/+Hi] segments from the inventory it was eliminated early. In tableau (3) the final sub-optimal candidate which represented a case of total harmony, was eliminated by the markedness constraints restricting the occurrence of a particular value for ATR. In this case the optimal candidate needed to be a lexical item in which the mid-vowels had reverted to the default form. In tableau (4) however the sub-optimal candidate with total harmony was eliminated by a different markedness constraint. In this situation the attested form reflected a case of total [+ATR] harmony due to vowel raising, therefore the default form is not necessary. The defining constraint in this case was the additional markedness constraint which eliminated the candidate with total [-ATR] harmony. Restricting the [-ATR/+Hi] segment allowed for the elimination of the candidate with total harmony for the wrong ATR value.

5.5.2. Partial harmony

Instances of partial ATR harmony are derived from lexical items with a combination of vowels with different height values. This includes constructions containing a combination of mid and high vowels, instances in which the high vowel does not trigger raising on all mid-vowels due to its position in the word. Therefore, some mid-vowels assimilate for the feature [+ATR] while the rest revert to the default form. There also instances in which the high vowel is incapable of triggering raising on any of the mid-vowels in the words, in which case all of the mid-vowels will revert to the default [-ATR] form resulting in a sequential disagreement for ATR. Lastly there are constructions which include both mid-vowels, high vowels and low vowels. In these instances there will always be sequential disagreement because high and low vowels do not share harmonic counter parts for ATR and therefore cannot alternate under any circumstance. In most cases of this nature the low vowel will serve as a blocker of any vowel raising, ensuring that the mid-vowels occur in the [-ATR] form. If the high vowels are still able to trigger raising on the mid-vowels within the construction this would still result in a case of partial harmony due to the occurrence of the low vowel. Every case in which harmony cannot be achieved throughout the entire construction may either be categorised as an instance of transparency or opacity. Both of these categories encompass cases of partial harmony, however they prefer a particular type of sequential agreement.

5.5.3. Transparency

In certain instances partial harmony includes some adjacent featural agreement between segments, but focuses on the total number of segments with a particular shared feature throughout the word. In these cases there may be an intervening vowel which disrupts harmony, however harmony is allowed to continue after this interference. In the harmony system of isiXhosa, the only way in which harmony may continue after being disrupted is if it is restarted by a second trigger, or if the initial pattern is [-ATR] and all high vowel spreading is blocked. In instances of transparent harmony, the system used to analyze the pattern must favor total harmony above harmony between adjacent segments. Hence, different forms of the same harmony constraints must be included, and these forms must differ in their reach and the segments they recognize. We therefore have local harmony constraints which favor adjacent segments, and distance constraints which push for total harmony favoring candidates with the highest number of agreeing segments.

(5) Partial Harmony: Constructions exhibiting transparency

/ebebhadlile/ 'he was caused to become angry'

/ebebhadlile/	*+ATR /+Lo	*-ATR /+Hi	*+ATR	*+ATR∞-ATR	*-ATR∞+ATR
ebebhədlile	*!		*****		
ebebhədlilɛ	*!		****	****	
ebebhədlɪlɛ	*!	*	***	******	
ebebhadlɪlɛ		*!	**	******	
ebɛbhadlɪlɛ		*!	*	*****	
ɛbɛbhadlɪlɛ		*!			
ɛbɛbhadlɪle		*!	*		****
ɛbɛbhadlile			**!		******!
ɛbɛbhədlile	*!		***		******
ɛbebhədlile	*!		****		****
ebɛbhədlile	*!		****	*	***
ebebhadlile			****!	**	**
ebebhədlɪle	*!	*	****	***	*
ebɛbhadlile			***!	**	****
ebebhadlɪle		*!	***	****	**
ɛbebhadlɪlɛ		*!	*	***	*
ɛbɛbhədlɪlɛ	*!	*	*	***	**
ɛbɛbhadlilɛ ☺			*	*	***
ɛbebhədlɪlɛ	*!	*	**	****	**
ɛbɛbhədlilɛ	*!		**	**	****
ebɛbhədlɪle	*!	*	***	**	*

ɛbebhadlilɛ			**!	**	***
ɛbebhədlilɛ	*!		***	***	***
ebɛbhadlılɛ		*!	*	****	
ebɛbhədlilɛ	*!		***	***	**
ɛbebhədlıle	*!	*	***	**	****
ɛbebhadlıle		*!	**	**	***
ebɛbhadlilɛ			**!	***	**
ebebhadlilɛ			***!	***	*
ɛbebhadlile			***!	*	*****
ɛbɛbhədlıle	*!	*	***	*	****
ebɛbhədlılɛ	*!	*	**	***	*

The partial harmony resulting from the analysis of the construction above, is a reflection of the capacity of the harmony system to derive forms which display transparency. More than half of the candidates are eliminated by the first markedness constraints restricting the occurrence of any segments which do not exist in the inventory. The third markedness constraint *[+ATR] then eliminates all candidates with mid-vowels in the [+ATR] form. As no raising occurs in the attested form of this construction, any optimal candidate must contain mid-vowels which occur only in the retracted form. The optimal candidate /ɛbɛbhadɬilɛ/ acquires one violation for the *[+ATR] markedness constraint, due to the non-alternating high vowel. As the high vowel is not in a position to cause any assimilation for [+ATR], defaulting to the [-ATR] form becomes a significant priority. If either of the harmony constraints were to be shifted above this markedness constraint in the ranking, the attested candidate would lose, as there are still candidates with no violations for these constraints. The attested candidate only has partial harmony and will therefore certainly attain violations for sequential prohibitions. The majority of segments in this case of transparent partial harmony share the feature value [-ATR]. To derive this kind of transparent harmony generally the distance harmony constraint *+ATR∞-ATR would eliminate candidates with too many [+ATR] segments. However, the limitation on the occurrence of [+ATR] segments is already achieved by the defaulting markedness constraint.

5.5.4. Opacity

Cases of opacity include instances of partial harmony in which substantial harmony occurs throughout a word, between adjacent segments. These constructions include a variety of segments, in which assimilation has occurred between specific adjacent segments. Once regressive raising has occurred due to a high vowel trigger, any preceding mid-vowels will assimilate for [+ATR], however the insertion of a non-alternating low vowel will discontinue this process and cause any further preceding mid-vowels to default to the [-ATR] form. Hence, these constructions will have both [+ATR] and [-ATR] harmony between adjacent segments. All instances of opaque partial harmony include constructions in which the high vowel is in the word final position preceding an alternating mid-vowel, then followed by an intervening low vowel which precedes the final word initial mid-vowel or mid-vowels. It may also be a construction in which the final mid-vowels have [-ATR] defaulting harmony, but are then preceded by a high vowel trigger, which then causes [+ATR] raising on any further preceding mi vowels. For these cases the ranking system must prioritise harmony between adjacent segments as oppose to total harmony within the entire word.

(6) Partial Harmony: Constructions exhibiting opacity

/esile/ 'donkey'

/ɛsilɛ/	*-ATR /+Hi	*-ATR/-Lo -C- +ATR/-Lo	*+ATR	*+ATR-C--ATR	*-ATR-C-+ATR	*-ATR
a. esile			***			
b. esilɛ ☺			**	**		*
c. esılɛ	*		*	**		*
d. ɛsılɛ	*					***
e. ɛsilɛ		*	*	*	*	**
f. esıle	*	*	**	*	*	*
g. ɛsile		*	**		**	*
h. ɛsıle	*	*	*		**	**

The majority of the sub-optimal candidates in the tableau above are eliminated by the initial markedness constraints prohibiting segments which do not exist in the inventory. There are two candidates not eliminated by the initial markedness constraints, however these candidates have not undergone the effects of high vowel raising. Raising must occur on any mid-vowel preceding a high vowel trigger. This [+ATR] assimilation is enforced by the newest harmony constraint placing requirements on both the ATR and height value of the sequential segments. All candidates in which raising does not occur are therefore eliminated, as raising is prioritised when the alternating segment occurs adjacent to a trigger. The two candidates without violations for these constraints, include the attested candidate and a sub-optimal candidate in which the high vowel has successfully triggered raising on the preceding mid-vowel. However, due to the directionality of the spread rule, only the preceding mid-vowel should occur in the [+ATR] form. The final mid-vowel must default to the [-ATR] form, and therefore the final sub-optimal candidate is eliminated by the markedness constraint *[+ATR].

5.5.5. Note about partial harmony in isiXhosa

There are undoubtedly a variety of attested forms in isiXhosa which exhibit instances of partial harmony. The harmony cases presented above display both cases of transparency as well as opacity between segments, however is this a product of the harmony system itself or simply a consequence of the segment functions. The construction above /ɛbɛbhadłilɛ/ reflects an instance of transparency, this is labelled a case of transparency because [-ATR] harmony appears to dominate the segments throughout the word. The high vowel is the only [+ATR] segment, and is not able to force any raising onto the preceding mid-vowel segments. Without the presence of the low vowel between the high vowel and all preceding mid-vowels [+ATR] assimilation would have occurred, which therefore proves that the high vowel is not transparent in nature but is prohibited from affecting any other segments. This case of transparency hence appears to be a surface phenomenon and not something intrinsic to the harmony system. This type of partial harmony will therefore always be a consequence of the positioning of the non-alternating segments within a construction. However, the analysis of opacity in partial harmony using /esilɛ/ indicates a definite systematic form of partial harmony. The opacity of the high vowel is accurately characterised, as the function of this segment is to trigger raising. Due to the extended requirements on directionality, the structure of the harmony pattern ensures that any mid-vowels positioned following a high vowel will default to [-ATR]. Hence, the pressures of regressive raising and defaulting ensure cases of opacity in partial harmony.

5.5.6. No harmony

Due to the complexity and exigencies of the harmony system in isiXhosa there are also attested forms in which there is no featural agreement between any of the adjacent segments or within the word as a whole. These include constructions in which there are mainly non-alternating vowel segments such as the high vowel pair and the low vowel [a]. The positioning of these segments are not prime for any form of raising or assimilation of ATR. It is imperative in these instances that the ranking prioritise the correct defaulting form for any mid-vowel segments, as such cases do not require raising.

(7) No Harmony: Constructions with a variety of non-alternating segments

/afikane/ 'scented grass, used to make knecklaces'

/afikanɛ/	*-ATR /-Hi	*+ATR /+Lo	*+ATR	*+ATR∞-ATR	*-ATR∞+ATR	*-ATR
a. əfikəne		**	****			
b. əfikənɛ		**	***	***		*
c. əfikanɛ		*	**	****		**
d. əfɪkanɛ	*	*	*	***		***
e. afɪkanɛ	*					****
f. afɪkane	*		*		***	***
g. afɪkəne	*	*	**		****	**
h. afikəne		*	***		***	*
i. afɪkənɛ	*	*	*		**	***
j. afikənɛ		*	**	**	**	**
k. afikanɛ ☺			*	**	*	***
l. əfɪkane	*	*	**	**	**	**
m. əfɪkəne	*	**	***	*	**	*
n. əfikane		*	***	**	*	*
o. afikane			**	*	***	**
p. əfɪkənɛ	*	**	**	***	*	**

The construction above has a total of 4 vowel segments, this results in an over-all set of 16 possible candidates. 14 of these candidates are immediately eliminated by the first and second markedness constraint. All of these candidates are therefore intolerable due to the inclusion of segments which do not exist in the vowel inventory. What separates the final sub-optimal candidate and the attested form is a final defaulting mid-vowel. The sub-optimal candidate /afikane/ is hence eliminated by the markedness constraint which enforces the correct default *[+ATR]. To reiterate the initial comment, a case of total harmony relies heavily on the enforcement of the correct default form. In the tableau above it is clear that the first three constraints perform the most salient functions, while the final constraints are not necessary to achieve harmony in this situation.

5.6. Discussion

5.6.1. Discussion of adaptation to No-disagreement approach

The constraints presented above reflect the OT approach of harmony by prohibition. Harmony is therefore accomplished by placing violations on disharmony (Pulleyblank 2002). This includes the prohibition on adjacent segments with different values for the same feature, realised in the harmony/agreement constraint (*–ATR - +ATR). However, this approach also involves the application of markedness as well as faithfulness constraints. We therefore have the following markedness constraints (*+ATR/Lo) & (-ATR/Hi) which function to eliminate a situation in which harmony is achieved using vowels which do not exist in the inventory. There are also basic markedness constraints (*-ATR) & (*+ATR) which restrict the occurrence of a particular value for ATR. Included in the ranking above is one faithfulness constraint (Ident ATR) which functions to place a final limitation on the number of accepted ATR alternations. Through the prohibition of disharmony one is able to avoid a number of predetermined postulations, and thereby analyse the lexical representation without a range of preconceived ideas. The purpose of this particular ranking approach is to avoid a dependence on the correspondence between what is referred to as the underlying form and the surface form, in order to produce a consistent analysis of the harmony pattern. The majority of the constraints included in this ranking involve a prohibition on segments, either on their phonetic composition or the feature values of consecutive segments. Hence, the faithfulness constraint which is included is one of the lowest ranked constraints, and therefore does not generally play a role in the selection of the optimal candidate.

Although there are constraint rankings which account for all attested candidates separately, it is revealed that none of these rankings has the capacity to deal with the empirical issues of each harmony case simultaneously. Therefore, not one of the formulated rankings could serve as an effective analysis of the isiXhosa harmony pattern. Based on the RCD information tableau processed in OT workplace, the issue resides in a contradiction between the selection of the attested forms which only include mid-vowel such as /ndɔnɛ/ 'I did wrong'. Forms such as /ɛtabeni/ 'mountain', which reflect a case of partial harmony containing a high vowel trigger, which spreads [+ATR] to the preceding mid-vowel segment, a feature withheld from the initial mid-vowel due to the blocking function of the low vowel. Hence, we have an attested candidate in which the two initial segments share the feature value [-ATR], followed by two segments sharing the feature value [+ATR]. As well as forms such as /ɔmalumɛ/ 'uncle' which could be viewed as exhibiting zero harmony since the final mid-vowel is not preceding the high vowel trigger and does not share the feature value [+ATR]. Hence, we have an attested candidate in which the initial two vowels share the feature value [-ATR] followed by two more vowels which are [+ATR] and the final vowel [-ATR]. Each of the attested candidates above are selected by constraints at which other attested candidates fail. This presents an issue because the defining constraints for each candidate are not able to form a comprehensive ranking system.

Beginning with the harmony constraints *[+ATR-C--ATR] & *[+ATR∞-ATR]. Although these constraints favour the attested candidate in most cases, they do not serve as the deciding constraint until the case of the candidates presented above. In the case of /ɔmalumɛ/ these constraints favour the losing candidate /ɔmalume/ which has partial harmony, unlike the attested candidate which reflects zero harmony. However, because these constraints favour mid-vowel candidates with only [-ATR] segments, they are necessary to select /ndɔnɛ/ above /ndonɛ/. Therefore, a constraint is necessary for the selection of the attested candidate with only [-ATR] segments in the case of mid-vowel only words. The next constraint which elicits contradictory results is the long distance harmony constraint *[-ATR∞+ATR]. In most cases this constraint favours the losing candidate as it prefers when most segments in a word surface with the feature value [+ATR], or have total agreement for [-ATR]. Of course in the isiXhosa harmony pattern total harmony is not always possible due to the specific harmony trigger and directionality requirements, as well as the blocking function of the low vowel.

The contradiction within the constraint ranking can be presented in the form of two choices. The choice between the markedness constraints 8 & 9, is also the choice between a default option of either [+ATR] or [-ATR]. Hence, if the ranking is 8>>9 the default option would be [-ATR] and all attested candidates for mid-vowel only words such as /ndɔnɛ/ would be considered optimal. However, if the ranking is 9>>8 then the value [+ATR] would be the default option, which would favour attested candidates such as /ɛtabeni/ as the /e/ would be [+ATR] and not [-ATR] assimilating with the low vowel [a]. Neither of these rankings are viable however as both candidates require the opposite default feature in order to win. The second choice we are faced with is reflected in the contradictory effects of constraints 3 & 6 vs. 4 & 7. These includes the local and distance harmony constraints *[+ATR∞-ATR] versus the opposite harmony constraint *[-ATR∞+ATR] and the faithfulness constraint [IdentATR]. The choice therefore becomes a choice between harmony and faithfulness. If harmony is ranked above faithfulness the attested candidate /ɔmalumɛ/ will win in most cases because it has only one violation for the harmony constraint, however against a candidate like /ɔmalume/ the attested candidate will lose as the new candidate has no violations for the harmony constraint. The real test the candidates fail however is against the input forms. In order for the attested candidate to win in all cases they must all be similar to the input and favoured by the faithfulness constraints.

5.6.2. Harmony cases causing contradictory ranking

(8) Cases with only mid-vowels:

/ndenze/ ‘I made’

/ndɛnzɛ/	**1:* +ATR, +Lo**	**2:* -ATR, +Hi**	**5: (*-ATR-C-+ATR)**	**3: (*+ATR~-ATR)**	**6: (*+ATR-C--ATR)**	**4: (*-ATR~+ATR)**	**7: Ident-[ATR]**	**8: *+ATR**	**9: *-ATR**
a. ndenze							*	**	
b. ndenzɛ				*					
c. ndɛnze			*						
d. ndɛnzɛ ☺									**

In cases of mid-vowel only words such as ‘ndenze’ the resulting optimal candidate should be one with total [-ATR] harmony. The ranking above achieves this harmony but requires that the final [+ATR] markedness constraint be ranked above the [-ATR] markedness constraint. The partial harmony candidates are consecutively eliminated by the harmony constraints as expected. These constraints push for total harmony either regressively such as (*-ATR-C-+ATR) which eliminates candidate (8c) above, or progressively such as (*+ATR∞-ATR) which eliminates candidate (8b). The remaining candidates which do not violate any sequential prohibitions both reflect instances of total harmony. The deciding constraint in the tableau above is the faithfulness constraint (Ident [ATR]), however an analysis dependent on faithfulness will not be sufficient, as the possible input form will not always match the attested form. As the input form consists of all possible candidates, there will always be cases in which the attested form will have segments which do not match those in the input form due to the pressure of the harmony drivers. Therefore, ignoring the tie to faithfulness in the analysis above the deciding constraint becomes the markedness constraint (*+ATR), which ensures that the default form is [-ATR]. In the analysis above it is therefore imperative that the markedness constraint (*+ATR) be ranked above (*-ATR).

(9) Cases of blocking and raising:

/etabeni/ ‘mountain’

/etabɛnɪ/	**1:* +ATR, +Lo**	**2:* -ATR, +Hi**	**5: (*-ATR-C-+ATR)**	**3: (*+ATR~-ATR)**	**6: (*+ATR-C--ATR)**	**4: (*-ATR~+ATR)**	**7: Ident-[ATR]**	**8: *+ATR**	**9: *-ATR**
a. ɛtabeni			*			****	***	**	**
b. ɛtabɛni ☺			*			***	**	*	***
c. ɛtəbɛni	*								
d. ɛtabɛnɪ		*							

In this case the attested candidate should be one of partial harmony, with both blocking and raising achieved. Blocking is achieved by the harmony constraints driving regressive spreading. The two harmony constraints (*+ATR-C--ATR) & (*+ATR∞-ATR) ensure that no [+ATR] segments precede [-ATR] segments, hence ensuring that the all mid-vowels preceding the low vowel [a] occur in the [-ATR] form. Neither of the candidates above violate these blocking harmony constraints, because both candidates have achieved blocking.

The remaining harmony constraints are those which drive the progressive spreading of [+ATR], these include constraints 5 & 4. The current ranking results in the incorrect optimal candidate, however because these candidates equally violate constraint 5 no shift in the ranking of these harmony constraints will result in the correct optimal candidate. Constraint 4, which drives harmony by forcing raising, eliminates the attested form as it forces raising beyond the requirements of the harmony system. It appears that no matter what the ranking of these harmony constraints both blocking and raising cannot be achieved as in candidate (9a). Ignoring the effects of faithfulness, as it has already been established that the input form and attested forms may be completely different, the only shift in ranking which would elicit the correct candidate is to rank markedness constraint 9 above 8. This is not a viable solution however, as the analysis in the previous tableau demonstrates that the markedness constraint restricting the occurrence of [+ATR] must be ranked above (*-ATR) in order to elicit the correct default form.

(10) Blocking as well as defaulting forms:
/omalume/ 'uncle'

/ɔmalumɛ/	**1:* +ATR, +Lo**	**2:* -ATR, +Hi**	**5: (*-ATR-C-+ATR)**	**3: (*+ATR~-ATR)**	**6: (*+ATR-C--ATR)**	**4: (*-ATR~+ATR)**	**7: Ident-[ATR]**	**8: *+ATR**	**9: *-ATR**
a. ɔmalumɛ			*	*	*	**		*	***
b. omalumɛ			*	***	**	*	*	**	**
c. ɔmalume ☺			*			**	*	**	**
d. ɔmalʊme		*							
e. ɔmǝlume	*								

In the tableau above based on the equal violations for each candidate on the first constraint, it is clear that they equally violate the sequential prohibition restricting the occurrence of [-ATR] segments preceding [+ATR] segments as raising does not occur. Blocking is achieved in candidates (10a) & (10c) as they have minimal violations for the regressive harmony constraint, which immediately eliminates the candidate without blocking. However, any ranking of the harmony constraints would still be unable to elicit the attested form and correct optimal candidate. In this case the constraints ensuring blocking require too much regressive harmony, thereby penalising the candidates with a final mid-vowel reverting to the default form. Excluding any reliance on faithfulness, the only ranking shift which would result in the correct optimal candidate would be to rank the defaulting markedness constraint above all harmony constraints. Cases such as ‘etabeni’ however which require both blocking and raising would not be harmonised accurately.

These constraints are clearly effective to some degree as they do exponentially limit the number of possible candidates. However, the application of these constraints alone cannot derive the harmony system found in isiXhosa. Depending on the ranking of the harmony constraints, harmony is either excessively enforced regressively or progressively. The pattern requires a nuanced compromise to achieve harmony, in which raising, blocking as well as defaulting may occur when necessary.

5.7. Discussion of analysis

The adapted approach to harmony by no-disagreement has proven effective in characterizing the harmony pattern found in isiXhosa. Each type of harmony reported to occur in the attested forms is addressed, and accounted for by the constraint ranking. Cases of total harmony are derived either by the harmony constraint driving regressive spreading, or by the markedness constraint forcing mid-variants to default to the [-ATR] form. More complex forms such as those representing instances of partial harmony are accounted for in similar ways. Generally the optimal candidate would be derived using the harmony constraints, simply editing the distance of the sequential prohibition. Hence, adjacent assimilation would be prioritised in cases of opacity, while total harmony across the lexical item would be prioritised to account for transparency. The analysis above has proven however with the markedness constraint *[+ATR] ranked above and below the appropriate constraints, and the adapted harmony constraint corresponding only between non-low vowels, the correct form of partial harmony is consistently derived. By ensuring raising is driven in the appropriate phonological contexts, all necessary [+ATR] assimilation is forced to occur. Those instances in which raising should not occur are then accounted for by the defaulting markedness constraint. In the overall ranking presented above, the highest ranked constraints are the markedness constraints prohibiting any segments which do not exist in the inventory.

Following this is the harmony constraint driving regressive harmony between non-low segments, and the most critically ranked constraint forcing the correct default form. Based on the conclusions drawn above however, the ATR harmony pattern in isiXhosa is driven by [+ATR] assimilation as a result of raising, or defaulting to the [-ATR] if raising is not required.

6. Acoustic Analysis

6.1. Introduction:

The following acoustic analysis is crucial to the process of answering the first and most basic research question posed in this paper; which features are relevant to the isiXhosa vowel harmony system? Previous literature has presented advanced tongue root/ATR as the definitive harmonic feature (Jokweni and Thipa 1996). This generalisation therefore serves as the preliminary hypothesis which is developed and strengthened by the introduction of acoustic data. The evidence produced from the results of this acoustic analysis offer an innovative explanation of the ATR agreement between the vowels in isiXhosa. The resulting data points to a visible discrepancy in both the height and horizontal position of the tongue in the production of the mid-vowel sets e, o & ɛ, ɔ. However, the substantial disparity appears mainly in the height differences between the mid-variant sets as well as the high and low vowels. Hence, what we find is a considerable difference in height between the vowel sets, which is consistent, and generally a less significant difference in the horizontal positioning of the tongue.

Based on the generalisations describing the vowel harmony system of isiXhosa, the mid-advanced variant should occur preceding a high vowel trigger, and the mid-retracted variant in every other situation (Jokweni and Thipa 1996). The use of acoustic software enables the investigation of mid-variant sets in every context, and allows for their categorisation based on the generalisations previously discussed. All mid-variants were extracted from a specific phonological context, and measured in isolation for their F1 and F2 values. If the preliminary generalisations are accurate, then the mid-variants preceding a high vowel should have a significantly lower value for F1 which would result in those vowels appearing higher within the vowel space. Furthermore, the retracted variants would then obtain higher F1 values and appear significantly lower in the vowel space (Ladefoged 2001). The F2 resonant values of the [+ATR] variants should systematically position them at the peripherals of the vowel space, in comparison to the [-ATR] variants (Fulop 1998). The comparison of these differing formant values are used to identify any shift of the articulators, and change in size of the pharyngeal cavity. Each of the questions posed below will be answered in the analysis ad comparison of particular sample groups. Hence, the effects raising and blocking are verified in the analysis of formant values extracted from segments in a particular phonological context. Segments indicative of raising effects are extracted from words in which they are positioned preceding a high vowel, while those indicative of the effects of blocking are position preceding a low vowel which must in turn be followed by a high vowels trigger. The comparisons of these segments extracted from different phonological contexts answer the question of where the ATR alternations are taking place and to what degree.

The aim for this chapter involves investigating and retrieving the acoustic value for each vowel used in isiXhosa. This data is then categorised to reflect the distribution of the different mid-variant sets in each phonetic context. The quantifiable data provided by this analysis allows one to objectively analyse the acoustic nature of the harmony pattern present in isiXhosa. This then leads us to a number of questions specifically asked and answered from an acoustic perspective:

(1) **Questions:**

1.1 Is there a real phonetic difference?

1.2 What kind of difference is it? Is it the harmonic feature [+/-ATR]?

1.3 Are the effects of the high and low vowels supported in the data?

1.4 Do the effects of harmony alternations work along a spectrum?

6.2. Acoustic properties of vowels

Vowels are generally viewed as a complete or single entity, therefore it is difficult to consider the separate parts of their production. The resonances of the vocal tract, also referred to as formants are impossible for human beings to hear separately. This is due to the contiguous nature of vowel segments (Reetz & Jongman 2009). Unlike consonants, due to the limited constriction in the production of a vowel and the high dependence on tongue and jaw movement in distinguishing the voice quality of a particular vowel, there is a higher level of contiguity between segments. Hence, in a CVC combination the vowel separating the consonants will require a complete reconfiguration of the vocal cavity, thereby resulting in the collapse of the initial consonant constriction. However, the combination VCV in which vowels are separated by a consonant, due to the localised articulation of the consonant, the free part of tongue and jaw are able to assume the shapes necessary for the production of the vowel segments. Hence, the movements are smooth and the tongue can easily shift from the shape of the initial vowel to the next (Gafos 2014). The application of acoustic software allows one to view the resonances and overtone pitches in the form of formants which constitute each vowel (Ladefoged 2001). The height and the front or backness of the vowel articulation are the factors which make up the most basic parameters of vowel description. All languages have vowels which contrast in terms of height and horizontal movement of the tongue in the articulatory process. These measures of constriction are best reflected through the differences between formant resonant values. Other features such as, nasality, lip position and pharyngeal space are considered to be superimposed on this fundamental vowel space (Lindau 1978).

The majority of world languages have developed a vowel system centred around the following three elementary vowels i,a,u, which are produced at the perimeter of the vowel space. Such even and distant distribution ensure the effective distinction between these vowels and therefore between different words. Vowels e & o are also considered quite easily distinguishable as they are produced at the intermediate spaces. Cases in which the distinction becomes more challenging to discern is within languages which make use of vowel systems as large as 7 or 9 vowels. In such systems vowels must be further distinguished by features such as length, height, roundness, or ATR, which may be identified by their formant frequencies (Ladefoged 2012). The African languages reported to have an ATR contrast vary in the size of their vowel inventory. It is quite rare to find a language with a symmetric inventory consisting of 10 vowels distinguished for ATR (Casali 2008). There are however numerous harmony systems reported to include between 9 & 7 vowel inventories. These include cases such as Yoruba (Archangeli & Pulleyblank 1989, Perkins 2005), Kinande (Archangeli & Pulleyblank 2002, Kimper 2011), Pulaar (Cole & Kisseberth 1994, Pulleyblank 2002, Sasa 2009), Nembe (Maduka 1988), as well as languages from the southern African Bantu language family, Setswana (Leroux 2008, Leroux 2013), Sesotho (Harris 1987, Khabanyane 1991), isiZulu (Harris 1987, Sibanda 2008) and finally isiXhosa (Jokweni & Thipa) on which the current research is focussed.

6.2.1. Co-articulatory effects of VCV sequences

Ideas and theories of co-articulation have emerged from the understanding of the concept of a segment. Many of the previous theories focus on the aspect of time, claiming that different kinds of acoustic signals are recorded, however, no distinct boundaries are established perpendicular to the time axis (Fowler 1980). Hence, no standard timeframe has been established against which to measure the beginning and end point of a segment. The concept of a temporally distinct segment is hence considered abstract in nature, as gestures are being produced simultaneously. The segment is theorised to exist only in the mind of the speaker and hearer. It is an idealised acoustic signal which cannot be physically produced but should simply be recognised as the target form. The process of articulation entails the continual shift of articulatory specifications as well as change in the global vocal tract. Hence, transitional sounds are consistently occurring between the realisations of any target forms. The only way to avoid these effects is if the sound is produced continuously in complete isolation to minimise any articulatory effect. Co-articulation should not be understood as the result of overlapping segments produced concurrently. Instead it is the alteration of the target form as a result of its context. Hence, the context is modified to accommodate the following gesture and it is this shift of the articulatory space that alters the production of the target form. Effects of co-articulation between VCV sequences are reported to extend only across adjacent segments and are highly dependent on the phonetic make-up of the preceding phoneme, evidence of which has been proposed in SiSwati (Malambe 2015). Previous analysis of this co-articulatory effect report that the intervocalic consonant constrains the articulatory movements in their timing. Furthermore anticipatory modifications for the second vowel are always initiated during the closure of the preceding consonant (Gay 1977). Work by Ohman (1966) reports that the co-articulatory relationship between vowels in a VCV combination may be described as diphthongal, the intervocalic consonant is basically superimposed on the acoustic signal of the vowels. Hence, articulation of the second vowel begins in during the closure of the initial vowel and is in fact independent of the consonant. This is however challenged by findings presented in Gay (1974) which tracked the movement of genioglossus in the production of different VCV syllables. The findings demonstrated separate muscle pulses in the production the vowel even when produced in symmetrical syllables. Furthermore continuous activity was recorded in the production of the consonant. The interpretation is therefore that there is a level of discontinuity in the movement between the vowels within the sequence.

6.2.2. Advanced tongue root or Extended pharynx

The feature Advanced tongue root (ATR) is most commonly recorded across languages in Sub-Saharan Africa. In its most straight forward instances the study of ATR includes vowel inventories of 10 vowel segments, with harmonic counterparts distinguished only by the feature ATR (Casali 2008). These languages include varieties of Sub-Saharan Africa, and many West African Niger- Congo languages. The feature ATR is referred to in various ways, however at its core it denotes the size of the pharynx in the process of articulation. The size of the pharynx is altered by the process of shifting the tongue root further forward and as a result higher, or retracting the root backward and thereby lower (Lindau 1978). The acoustic effects of this feature have been reported on but not yet studied in isiXhosa. This phenomenon results in a form of vowel harmony comprised of 2 sets of segments, categorised according to the size of the pharynx during the process of articulation. One set includes vowels produced with a comparatively larger pharynx [+ATR], the other set will include vowels with a pharynx that is comparatively smaller [-ATR]. In the case of isiXhosa however these harmonic sets are limited to the mid-vowels in the inventory. The only vowels with ATR counterparts are the mid-vowels, which means that only one set of vowels has been added to the inventory (Guion, Post & Payne 2004). Therefore, unlike the more straightforward instances of ATR harmony, isiXhosa has an asymmetrical 7 vowel inventory. Hence, the sets distinguished by pharynx size are not equal in the number of vowel segments included. Set 1 will include 4 vowel segments, the 2 high vowels and 2 mid-vowels [+ATR]. The second set then includes only 3 vowel segments, the second harmonic set of mid-vowels [-ATR] as well as the low vowel. The harmony system should then select vowels from the same set in situations in which harmony is required (Bakovic 2000).

Many harmony patterns of this type in African languages, have been proposed to involve a variety of distinguishing features. In many cases the harmonizing feature is defined as either; Breathy, Raised Height, Covered or Tense. However, radiographic data has proven that in these harmony types the harmony pattern is phonetically controlled by the movement of the tongue root (Ladefoged 1964, Retord 1972, Painter 1973, Lindau 1974). Furthermore, it is also demonstrated that the back pharyngeal wall and the vertical shift of the larynx work in conjunction with the tongue root mechanism, in order to achieve variation in pharynx size (Lindblom & Sundberg 1971). Hence, although referred to by different labels the harmonic feature in these cases may be definitively described as ATR. Taking this detailed definition into account an ultrasound analysis of this phenomenon would form a significant contribution to establishing a definitive description of this feature. Based on the explanation above there is a distinct articulatory-acoustic relationship involved in this harmony phenomenon. The discrepancy in pharynx size which has been identified as the difference between the relevant harmonic sets, will affect the frequency of the formants due to the reconfiguration of the pharyngeal cavity. The frequency of the first formant is therefore not only an indication of the height, but should reflect acoustic similarity or even the merging of vowels of distinctly different heights between the two sets. Hence, the mid-vowel from set 1 with the expanded pharynx should share a similar F1 value to the high vowel from set 2 with a reduced pharynx size. This similarity in the acoustic effect of the variation of the highest point of the tongue and the size of the pharynx, may serve to explain the mischaracterisation of this type of harmony, which has generally been described as a height centred phenomenon (Lindau 1978). The horizontal position of the tongue root appears to be a consequent variable of the change in pharynx size. Hence, the value of second formant should also serve as an indication of this mechanism.

Particular studies have chosen to emphasise duration as an indicative feature of ATR. The distinction between tense & lax vowels in English display a significant degree of discrepancy in length. The lax vowels are reported to have a shorter durations than that of the tense vowels. A distinction between durations was however not found to be significant in a comparison of [+ATR] & [-ATR] variants in a study of Maa conducted in 2004 (Starwalt 2008). Further analysis of differences in duration between ATR variants in Akan and Yoruba exhibit little to no significant discrepancy. Another acoustic feature considered useful in the analysis of ATR variant is the bandwidth of formants. As the general bandwidth correlates to the formant frequency, variants with a higher frequency are expected to possess a wider bandwidth. What was found is a greater correlation between the F1 frequencies and B1 frequencies of the [-ATR] variants as opposed to the [+ATR] variants. Generally however the most indicative acoustic feature include the formant frequencies (Starwalt 2008). The feature advanced tongue root has been reported to share the same basic phonetic mechanism as the tense/lax feature. The tongue root being the most active contributor to the change in pharyngeal cavity. The most clear and consistent way of testing this alteration within the vicinity of the tongue root, within the vocal tract, is to measure to formant frequencies.

6.3. Method

The entire analysis described below was performed using novel recorded data see Appendix chapter three. Words and phrases include both extracts from the greater isiXhosa dictionary (Tshabe, Shoba, Mini, Pahl, Pienaar & Ndungane 2006) as well as improvised utterances produced by the speaker. The greater dictionary of isiXhosa proved an invaluable resource in finding words with specific vowel and morpheme combinations. This resource enabled the search for words which only have mid-vowel roots, and the relevant affixes which could be attached to a specific lexeme. Majority of the words elicited were extracted from the dictionary, while the improvised utterances came directly from the speaker as an alternate option. All data was collected in a total of 6 sessions ranging between 2-3 hours of elicitation per session. The recordings were produced in a sound lab using a digital recording device, specifically a zoom h4n recorder. The microphone was a Nady HM-10 mic, with XM-10 phantom power. All data was elicited from one speaker, a 22 year-old female. Born and raised in King Williams town, she speaks the dialect isiRharhabe which is a dialect spoken in the former Ciskei. As the greater isiXhosa dictionary contains various words which are quite formal and even archaic, the elicitation involved her reading off of the word lists numerous times. In many instances however the colloquial utterances were elicited as natural responses.

6.3.1. Praat software & acoustic analysis

Unlike consonants vowels cannot be most accurately described by their place and process of articulation, but rather according to their acoustic properties. As vowels are produced with a relatively free flowing airstream the resulting sound therefore generates a strong acoustic signal. The quality of a particular vowel is determined by the shape of the vowel tract which in turn defines the value of the formant frequencies. The formant frequencies will therefore be modified by any shift of articulators or change in shape (Reetz & Jongman 2009). To perform an acoustic analysis of vowels one requires recorded data which can be processed in either phonetic or acoustic software such as PRAAT, which is the software used in this particular analysis. This graphical representation consists of individual wave forms, as well as multiple formants denoting overtone pitches and vocal-chord resonances (Boersma & Weenink 2005). This creates a clear image of the vowel space and allows one to make deductions about the position and featural characteristics of each vowel, but more importantly between two vowels. Thereby enabling a clear distinction between the vowels e,o & ɛ,ɔ in isiXhosa which hold different values for the feature ATR, but are not contrastive in nature and therefore are not represented separately in the inventory.

6.3.2. Data & elicitation process

273 recorded words and phrases make up the collected data for this experiment. Each word ranging between 2-5 vowels, resulting in an average of three vowels per word. This data consists of a selected number of individual words and in a few cases phrases containing different realisations of the mid-vowels ɛ & ɔ e.g. Okwe baloni igqabhuka ‘To burst like a balloon’. This utterance includes an instance of mid-vowel defaulting [-ATR] in ‘<u>o</u>kw<u>e</u>’, as well as mid-vowel raising [+ATR] in ‘bal<u>o</u>ni’. The purpose of this selection strategy is to ensure that the featural properties of the isiXhosa mid-vowels are studied in multiple conditions, to account for the influence of any preceding or following vowels. The acoustic signal of an advanced mid-variant must therefore be analysed in the phonological context in

which the mid-vowel precedes a high vowel trigger, while the defaulting mid-variant can only be identified in a phonological context including only mid-vowels, or with minimal high or low vowel influence. Due to effects of co-articulation and locality in harmony the extent of vowel raising or spread of [+ATR] must be established. Based on the generalisations discussed in chapter three, there are a number of circumstances and contingencies which may force or limit the realisation of a specific mid-variant form. Hence, the recorded constructions are divided into different categories generating 6 distinct vowel combinations. Illustrated below are the types of combinations investigated. All of the collected data can be categorised into one of the 6 groups below.

(1) Vowel combinations:

Raising	[E/O-i/u]	iqhoni → [iqhoni]	‘petticoat of various colours’
Directionality	[E/O-i/u-E/O]	ibhekile → [ibhekilɛ]	‘tin/can’
Extent of [+ATR] spread	[E/O-i/u-a]	phetula → [phetula]	‘turn about’
Blocking	[E/O-a-i/u]	ebebhadlile → [ɛbɛbhadlilɛ]	‘she/he became angry’
Extent of [-ATR] retraction	[E/O-a]	ndiyeva → [ndiyɛva]	‘I hear’
Defaulting	[E/O]	uqengehe → [uqɛngɛhɛ]	‘cheat/trickster’

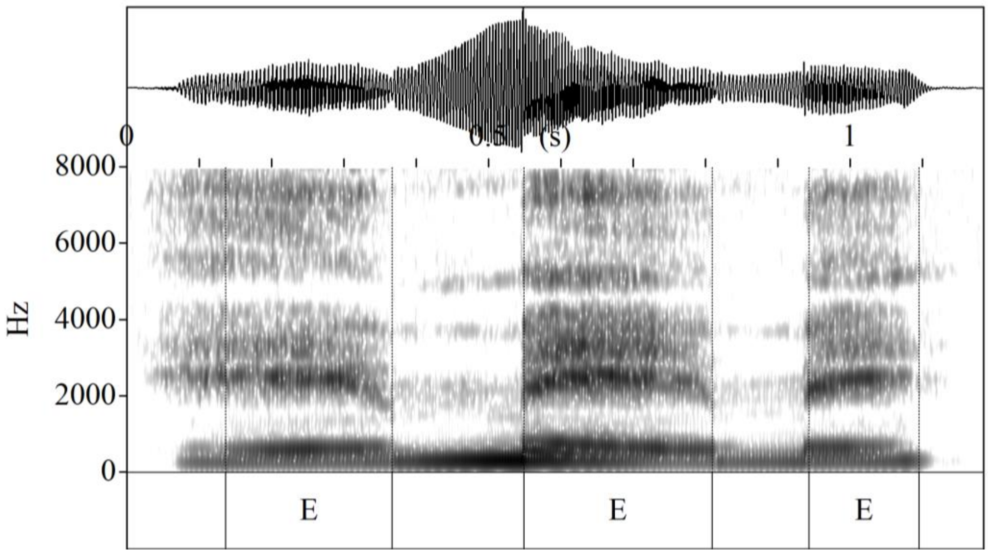
The first combination [E/O-i/u] considers basic raising contexts with a mid-vowel preceding a high vowel. According to the concept of vowel raising, the mid-variant in this instance should surface as the advanced form retaining the feature [+ATR]. The following high vowel should trigger vowel raising and [+ATR] assimilation in the preceding mid-vowel. The raising generalisation is therefore directly tested by the mid-variants which appear in these constructions. The combination [E/O-i-E/O] investigates the implication of a high vowel both following and preceding a mid-vowel. This is a test of directionality, to observe whether the position of the mid-vowel in relation to the high vowel trigger is of any relevance. Hence, will the high vowel cause the advancement of both the preceding and following mid-vowel, or will the assimilation of [+ATR] work regressively in alignment with the proposed generalisations.

The next combination [E/O-a-i/u] serves to distinguish the role of the low vowel [a] in disrupting harmony. As previously stated the low vowel [a] is the only vowel without a harmonic counterpart for both ATR and backness and is therefore considered a neutral vowel (Van Der Hulst & Van De Weijer 1995). Although neutral in its existence, this vowel is not transparent in its role within the harmony system. The generalisation states that [a] will cause disharmony in the process of blocking any raising triggered by the high vowel it precedes. The preceding mid-vowel is therefore prohibited from undergoing assimilation for [+ATR], which results in disharmony within the relevant domain. The fourth combination [E/O-i/u-a] demonstrates that [a] may only disrupt the process of vowel raising when it is positioned preceding the trigger and following the mid-vowel/s. Hence, the specific directionality of raising limits the blocking capacity of [a]. If [a] does not occur on the appropriate side of the trigger it will not stop [+ATR] assimilation. The final combinations include [E/O-a] & [E/O]; the first tests the implication of a mid-vowel before the low vowel [a]. Due to co-articulatory effects the retracted mid-variant in close proximity to the low vowel should show signs of assimilation and possibly occur further retracted than the default form. The final combination considers words in which there are only mid-vowels or words which contain only mid-vowels in their roots. The mid-vowel form situated neither in close proximity to the low vowel or high vowel should indicate the general acoustic shape and value of the default form.

6.3.3. Vowel annotation & extraction

In order for these combinatorial categories to be useful, each vowel must be annotated as the predicted variant. Therefore, using the symbols E & O to represent the [–ATR] variants and e & o to represent the [+ATR] variants, each mid-vowel is annotated based on its position relevant to the other vowels. In a case resembling the first combination above [E/O-i/u] the mid-vowel/s are annotated as their [+ATR] variants e/o. This [+ATR] annotation is due to the pattern of [+ATR] regressive spreading. In cases resembling combinations such as [E/O-a-i/u], [E/O-a], [E/O] the mid-vowel/s are annotated as their respective [-ATR] harmonic counterparts E/O. These mid-vowels are therefore annotated as retracted, first due to [+ATR] blocking performed by the low vowel, while the other two cases occur as retracted due to the defaulting structure. This form of annotation predicting the realisation of each variant within a particular phonological context is useful in determining the accurateness of each generalisation. Therefore, if the generalisation presented in Jokweni & Thipa (1996) states that all high vowels act as a trigger and cause regressive spreading of the feature [+ATR], then this should be reflected in the results of the annotated data. All mid-vowels preceding a high vowel should therefore accurately be annotated as [+ATR]. The other side of the token states that all mid-vowels which occur without a trigger should be annotated as the retracted form [-ATR].

An investigation into the acoustic features of each vowel is necessary in order to answer the research questions posed above. Each vowel is therefore extracted and measured individually, this includes the high vowels [i] & [u], the mid-variant sets [e], [o] & [ɛ], [ɔ] and finally the low vowel [a]. For the most accurate frequency values the vowel measurement must be extracted from the centre of that particular acoustic signal. The mid-point was chosen to reduce the effects of segment transitions (Starwalt 2008). By selecting a central portion of the vowel effects of the onset and coda consonants as well as breathiness of any preceding consonant are greatly reduced. To ensure consistency and accurate measurement extraction a PRAAT script was used to perform this function. The process of identifying the start and end point of each vowel involves tracking the formant frequencies while listening to the recording. As can be seen in figure (2) below the beginning and end point of each vowel must be distinctly marked. It is also clear how the difference in F1 and F2 values contribute to identifying the moment of inertia between different vowels.



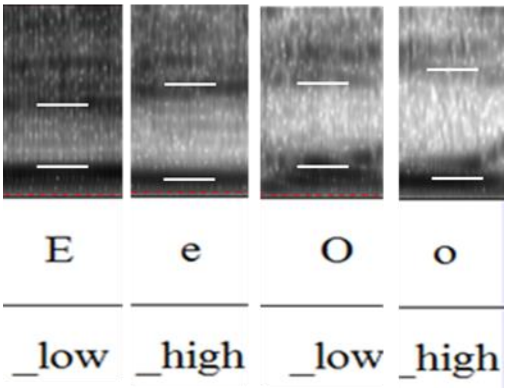
(2) Figure: The mid-point of a set of front mid-vowels. Hememe ‘you’re in trouble’.

The separated vowel sounds are then transcribed. The mid-vowels are further annotated for their ATR value. This annotation is completely dependent on the position of the mid-vowel within one of the particular vowel combinations above. Once the relevant sample is annotated, the PRAAT script is used to extract the measurements of the first and second formants at the midpoint of each vowel. As indicated above the first and second formants are the most relevant values for depicting the height and position of the tongue in vowel production. Once plotted, the vowels have been compared and contrasted. With a clear image of the vowel space deductions were made about the position and featural characteristics of each vowel, but more importantly between two vowels. Thereby enabling a clear distinction between the vowel sets e, o & ɛ, ɔ in isiXhosa, which are expected to retain different values for the feature ATR but are not considered separate phonemes.

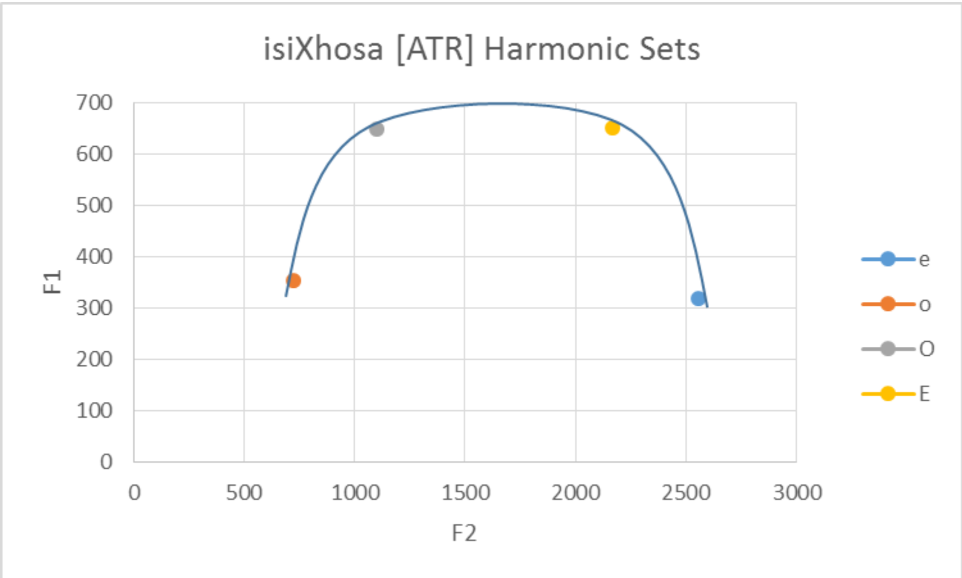
6.4. Relevant formant values

The relevant formants for the purposes of this analysis include the first and second formants, which provide an indication of the height and horizontal position of the vowel. According to previous literature the average measurements of the first and second formants offer the most beneficial representation of the vowel (Halle 1969, Leroux 2008, Lindau 1978, Reetz & Jongman 2009, Startwalt 2008). Referring back to the feature ATR in isiXhosa, if accurately described in the literature the tongue position should differ between the respective mid-vowel sets. The [+ATR] mid-vowels e & o are produced with the tongue root shifted forward resulting in variants with more closure, which as a result of the horizontal shift are also located higher within the vowel space (Ladefoged & Disner 2012). These changes in the shape of the vocal tract are then reflected in the frequency values of the relevant formants. As [+ATR] variants are higher and more peripheral, this results in different F1 & F2 values than [-ATR] variants, which occur lower and more centred.

In order to accurately interpret the relevant formant values for each vowel it is important to understand the structural nature of the charted vowels. Presented in Figure (3) below are the four different variants of the isiXhosa mid-vowels orthographic e & o. Based on this representation the higher the pitch the lower the vowel, hence the F1 value of segments E & O are higher than that of e & o. This variation in height is clearly displayed in the density and physical height of the F1 frequencies in relation to each other. In agreement with the interpretation above, in the process of plotting these measurements the scale for the first formant should be displayed along the vertical axis to produce the most accurate representation of height see figure (4). In the chart below the [-ATR] mid-variants are plotted higher due to their high F1 frequency values. The higher segments within the vowel space are however the [+ATR] mid-variants which share low values for F1. Hence, a more transparent display of this data would reverse the order of the F1 values along the y axis. It is also clear that the value of the second formant for the mid-variant /e/ is higher than that of its harmonic counterpart /E/. This serves as an indication of the forwardness or backness of the tongue position in producing the vowel. Therefore, the variant /e/ is not only produced at a higher level than its harmonic counterpart but is also produced further to the front of the mouth. Although this serves as a useful indication of the position of the tongue in the production of these vowels, one must acknowledge that the second formant also considers the roundness of a vowel (Ladefoged & Disner 2012). Hence, the combination of these formant values serve as a better representation of this particular feature as opposed to contrasting the individual F1 or F2 values of each segment.



(3) Figure: Comparison of extracted formant frequencies.



(4) Figure: The scatterplot & line graph above reflect the F1 and F2 measurements of the mid-variants in the following constructions; e- 'entlekeleni, o- 'iastronomi', E- 'iqengele', O- 'isondo'.

The extracted measurements for each vowel are then plotted in R studio, producing a range of vowel charts. The distinctive feature of these vowel charts are the visible discrepancies between the ATR mid-vowel variants. Each variant is charted separately, even those which have been analysed as non-phonemic. The F1 and F2 values extracted from the mid-point of each vowel on PRAAT were plotted onto vowel charts in the R studio programme. Using a command script the F1 values are plotted along the y-axis, while the F2 are plotted along x-axis (Baayen 2008). The resulting charts are presented in the analysis section below. Each variant is charted separately and represented within the category of vowel combinations in which they occur. This enables the direct comparison of mid-vowel variants which occur as a result of raising, only influenced by the following high vowel [E/O-i/u] vs. those which occur preceding a low vowel and then a high vowel [E/O-a-i/u]. Thereby limiting the variables and testing the results of raising and low vowel blocking with minimal interference. A number of tokens were however excluded from the analysis in R Studio, these include all the word final tokens except for those testing directionality. As word-final vowels are normally reduced and often devoiced in Xhosa these tokens were excluded as an extended measure to avoid any anomalies skewing the data. The word final vowels are not fully articulated and will therefore not reflect the complete acoustic nature of those sounds. Furthermore even with these efforts there remain a number of outliers, tokens which did not pattern exactly as expected. However, the resonances of these tokens appeared to be unmeasurable and so were not accurately recognised by the formant tracker.

6.5. Statistical analysis

A statistical analysis of the numerical data is imperative in order to determine whether the relevant phenomenon is at all significant. Although the extracted measurements and vowel plots are extensively indicative of visible differences between the relevant segments, the differences found and commented on must be proven to be statistically significant. Hence, a statistical analysis of the data is presented to provide an accurate calculation of differences between samples and variables, as well as insight into their interpretation (Chandler 1987). Previous studies of this type have applied two-tailed t-tests to calculate the significance between the mean value of different variables (Malambe 2015). For the purposes of this analysis a form of descriptive statistical analysis is applied to the data. Therefore, specific calculations have been performed on a sample data set, the results of which are then contextually interpreted. In cases in which a significant difference has been calculated between data sets, the p-value will be smaller than the level of confidence. The level of confidence may theoretically be realised as any value however the general tendency is 0,05 which is the value used in this analysis. The null hypothesis which assumes no significant contrast between two data sets/vowels is solidified by a p-value greater than the level of confidence but is rejected if the p-value is smaller (Baayen 2008).

6.5.1. Two-way ANOVA with replication

The collected data for this analysis includes two formant measurements of the five underlying vowels in isiXhosa, as well as a set of [+ATR] mid-variants. The samples included in each statistical analysis consist of two groups, two separately annotated vowels, each with two independent variables. This type of data set therefore requires a particular statistical formula capable of analyzing two variables across two separate groups. The appropriate calculation for this analysis is a two-way ANOVA with replication. This calculation performs F-tests on each variable within a group, across groups, as well as between the mean values of each variable within each group. The replication of the ANOVA calculation enables the comparison of two variables across each group. In the present analysis these results are presented in the form of the following ANOVA tableau (Bluman 2000).

(5) :

Mid-front variants	average		P-value	
	F1	F2	Sample	0,171
ɛ	435	2429,941	Interaction	0,024
e	401,235	2567,147		

The numeric values of two mid-variants are compared above, the retracted mid-variant [ɛ] and the advanced mid-variant [e]. This is one of the most important comparisons of this analysis, which is statically proven to share a significant relationship in the tableau above. The first two columns include the calculated average values of the first formant and the second formant for a particular vowel. The final column with numeric values include two different types of p-values. The sample p-value describes the relationship between the formant mean values of each vowel. This p-value therefore determines whether there is a significant difference between the accumulative values of the tokens of each variant. If the mean value of these formants are significantly different, then these particular phonetic segments are distinctly different. This is not the case with the mid-variants above as the sample p-value is greater than 0, 05. The significance of the relationship between these variants is not however solely determined by this comparative value. The final p-value in the tableau is the interaction p-value, this determines the significance of the interaction between the formant values and a particular mid-vowel variant. Therefore, if the p-value is smaller than the level of confidence as in the tableau above, this means that there is a significant level of co-dependency between the formant values of the tokens of a particular vowel and the vowel itself. This is an important relation as it places limitations on the range of formant values that constitute a particular ATR variant (Bluman 2000).

6.5.2. Sample size

The sample sizes required for a statistical analysis are not expected to extend beyond a minimum size, however the samples being compared must be equal in size. Therefore, the data set for one group cannot be larger than the data set of the group to which it is compared (Chandler 1987). The statistical analysis included in this investigation is conducted using data sets which vary in size. There are numerous constructions with mid-front variants in all the phonetic combinations analysed here, however there are far fewer constructions with mid-back variants. Therefore, the sample sizes of the mid-front variants are in certain cases much larger than those compared across the mid-back variants. As the data sets must be equal between the two groups being compared, one sample size may need to be reduced in order to be compared to another. Hence, there are numerous tokens of the default mid-back variant [ɔ]. However, when compared to the mid-back variant extracted from the combination [ɔ-a], the default data set must be reduced to account for the lack of mid-back variant tokens recorded in those particular constructions.

6.6. Acoustic results & interpretation

The resulting vowel plots produced in accordance with the investigation above provide new insight to the acoustic nature of the isiXhosa vowel system. The vowels have been plotted according to their F1 & F2 resonant values and therefore occupy an objectively demarcated and limited vowel space. Each chart reflects the values of mid-variants occurring within specific vowel combinations. The results of these isolated investigations of mid-variants within different phonetic contexts allow one to answer the questions in illustration (1) above, by evaluating the differences between the individual sets of mid-variants. In each instance there is a visible difference in the position of the mid-variant in comparison to the high and low vowels. The charts which include plots of both mid-variant sets display significant disparity between the advanced and retracted mid-variants.

6.6.1. What is the default mid-vowel form?

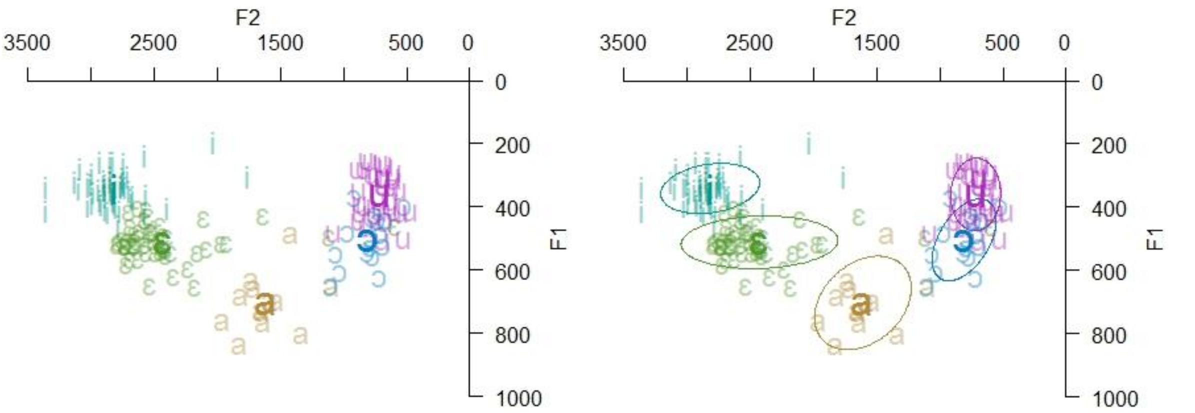
The formant resonant values plotted in the charts below are exclusively extracted from constructions in which the mid-vowels are unaffected by any high or low vowel influence.

(6) Default mid-variants:

a. iqengele ‘a swollen part of the body e.g. tumor’
b. iqokobhe ‘empty rigid container or shell’

The effects produced by high and low vowels have been observed to occur only regressively, hence any non-mid prefixes are entirely ignored. The ATR value of the mid-variants presented in the examples above, and plotted in the charts below therefore serve as an indication of the default ATR value, and the most accurate representation of the underlying mid-vowel set in the isiXhosa inventory.

Plotting vowel combination [ɛ/ɔ]:



(7) Figure: The vowel charts above include the vowels inherent to the isiXhosa vowel system. The mid-variants that are plotted above serve as the default forms, which appear in vowel combinations without the influence of either high or low vowels.
[i] = [+ATR] turquoise; [u] = [+ATR] dark pink; [e] = [+ATR] dark green; [o] = [+ATR] dark blue; [ɛ] = [-ATR] green; [ɔ] = [-ATR] purple; [a] = [-ATR] orange.

The vowel charts in figure (7) above include the plots of five isiXhosa vowels. The chart on the left displays only the vowels, while the chart on the right includes inserted ellipses. The ellipses encompassing each vowel space were calculated in R as a method of providing the location of the mean of each vowel. Calculated based on a presupposed confidence level and the covariance of tokens, the ellipses provide a type of summary of the distribution of the data (McCloy 2016). The vowels plotted in the chart above include only those from the vowel combination [ɛ/ɔ]. As presented in the examples above this combination includes constructions which either consist solely of mid-vowels, or have mid-vowel roots. The mid-variants measured therefore represent the default option, which reflects the surface realisation of mid-vowels in constructions without high or low vowel influence. What we find in the data represented above are five distinct vowels, which is what one would expect to see in a general quadrilateral vowel chart. The front high and mid-vowels are situated at different heights, with the mid-vowel appearing distinctly lower at just above 600 Hertz and the high vowel just above 400 Hertz. There is also a distinct difference in the F2 resonant values of these vowels, the front mid-vowel occurring further toward the back of the mouth than the high vowel, and the back mid-vowel positioned closer to the front. The ellipses which measure the trends of the separate vowel data points, attribute a distinct vowel space to each of the plotted segments. Each segment visibly occurs within a designated and limited space according to this statistical representation. Considering the sounds presented above are separate phonemes, this is exactly what one would expect to see in an analysis of their acoustic features.

6.6.1.1. Comparison of vowel averages

(8) IsiXhosa vowel formant averages

Vowel	F1	F2
[i]	341	2813
[u]	360	716
[ɛ]	509	2430
[ɔ]	499	799
[a]	702	1607

The initial set of average formant values in figure (8) have been calculated according to the measurements of each of the five vowels found in the isiXhosa inventory. The difference in F1 and F2 value averages of each of the high and mid-vowels is quite substantial. A difference of over 100 hertz appears consistently in the comparison of high and mid-vowels, while the value of the low vowel indicates a difference of as much as 400 hertz in comparison to the high vowel, and is at least 200 hertz higher than the mid-vowels. The average resonant values for the second formant also display consistent indications of alteration of the pharyngeal cavity. This is demonstrated by the difference between the front high and mid-vowel, and the back high and mid-vowel. Between the front vowels [i] & [ɛ], the mid-vowel has a lower average F2 value and would therefore occur slightly further back in the vowel space as demonstrated in figure (7) above. The back mid-vowel [ɔ] then has a slightly higher average F2 value than the high back vowel [u], with a difference of about 70 hertz. Hence, the mid-back vowel occurs slightly closer to the front of the vowel space than the high vowel, see figure (7). The mid-vowel variants presented have been annotated as [-ATR] due to the phonological context from which they were extracted. As they are considered [-ATR] variants in isiXhosa the averages have been contrasted with the formant averages of the retracted mid-variants reported below. This comparison serves to ascertain whether their F1 & F2 average resonant values are also indicative of a [-ATR] value.

Tableau adapted from Le roux & Le roux (2008): Setswana vowel averages

(9) Formant value averaged in Setswana

Vowel	F1	F2
[i]	329	1957
[u]	306	811
[ɛ]	507	1734
[ɔ]	503	921
[a]	758	1321

The vowel averages calculated for the Bantu language Setswana (Le roux & Le roux 2008), are within a similar range as those observed in isiXhosa. Although generally the average F2 values display a higher discrepancy between the vowels in isiXhosa compared to those in Setswana, overall the values are within the same ranges. The mid-vowels measured and calculated for isiXhosa include only those considered to be unaffected by any high or low vowel influence, hence these mid-vowels are considered to be the default choice and therefore inherent to the vowel inventory. The average values for Setswana however include measurements specifically for the retracted mid-vowels [ɛ] & [ɔ]. Hence, while the mid-vowel values displayed in figure (8) above include the default option for isiXhosa, those in figure (9) represent the retracted mid-vowels recorded in Setswana. The similarity between the F1 and F2 values for these mid-vowels therefore point to the retracted nature of the defaulting mid-vowels used in isiXhosa. This serves as further substantiation for the postulation of the [-ATR] form as the defaulting option and the unmarked choice.

Although Setswana has a total of seven phonemic vowels and isiXhosa only five, the assertion is not that the two data sets of the same segment are being compared. The average formant values of the retracted variants calculated by Le roux & Le roux (2008) serve as a type of benchmark against which to judge the retracted nature of the underlying mid-vowel in isiXhosa. While in this case there is a close correlation between the average values of these [-ATR] variants, these formant values are not always consistent cross-linguistically. Hence, the phonetics of one ATR language will not necessarily match that of another (Guion 2004).

6.6.2. Effects of raising [+ATR] spreading

The formant values plotted in the charts below are extracted from two separate sets of vowel combinations. The advanced mid-variant values below include only those which occur in the phonological context preceding a high vowel (10 a), and are therefore the prime examples of a raised mid-variant. These are then plotted against retracted mid-variants which have been extracted from phonological contexts in which the mid-vowel occurs uninfluenced or affected by either high or low vowels (10 b), this is therefore considered the default mid-vowel.

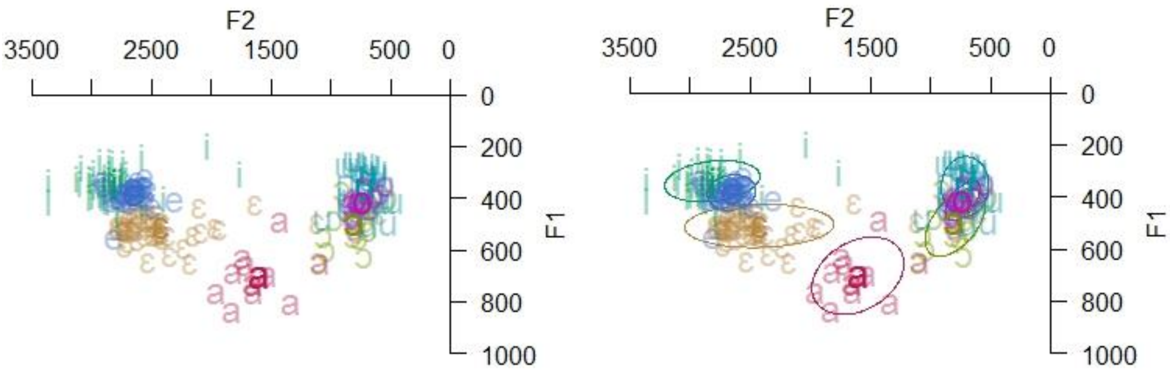
(10)

a. ibhedengu 'liar, thief' (Advanced mid-variant)

b. ndenze 'I made' (Default mid-variant)

The mapping of these particular mid-variant occurrences below is necessary to determine the extent of [+ATR] assimilation. Therefore, when a mid-vowel undergoes the effects of vowel raising, how high is it raised? Hence, what is the difference between the default mid-variant and the raised mid-variant?

Plotting vowel combinations [e/o-i/u] & [E/O]



(11) Figure: The vowel charts above includes two sets of vowel combinations. The [+ATR] mid-variants include those phonologically preceding a high vowel, and have hence assimilated for ATR. The [-ATR] mid-variants are the mid-variants inherent to the vowel system, which appear in vowel combinations without the influence of either high or low vowels.
[i] = [+ATR] turquoise; [u] = [+ATR] light blue; [e] = [+ATR] dark blue; [o] = [+ATR] purple; [ɛ] = [-ATR] brown; [ɔ] = [-ATR] light green; [a] = [-ATR] dark pink.

The segments mapped in the vowel charts above include a total of seven distinct vowels. The [+ATR] mid-variants [e/o] are the segments which most accurately represent the effects of vowel raising. These variants include those exclusively extracted from vowel combinations in which the mid-vowel occurs preceding a high vowel. The [-ATR] mid-variants plotted in these charts are extracted from a different combination without any high or low vowel influence, and are representative of the default mid-variant. The formant resonant values of these variants plotted together are useful in determining the extent of high vowel raising. Considering the resonant values of the defaulting mid-variants should provide insight to the shift in height and horizontal positioning of a raised mid-variant. The initial chart in figure (11) above illustrates visible differences in the positioning of both [+ATR] & [-ATR] mid-vowel tokens. The F1 values of the front [+ATR] mid-vowel tokens generally appear to be at or above 400 Hertz, while the [-ATR] tokens occur within a range of 400-600 Hertz. The F1 values of the back [+/-ATR] mid-vowels occur within the same ranges, however exhibiting less of a distinction than the front vowels. The height distinction between the mid-variant sets is therefore more prominent in the case of the front vowels, while still visible between the back vowels. The difference in F2 values is less distinctive and is better examined by considering the differences between the mean token values. The front mid-variant sets are distinguished by their mean token values being above 2500 Hertz [+ATR] or below [-ATR]. The mean tokens of the back mid-variant sets again display a lesser difference with the [+ATR] variant occurring closer to 500 Hertz than the [-ATR] variant, however both are positioned within the same range. It is clear from the charts above that the front mid-variants display a more prominent difference across all formant values. The F1 values for both front and back [+ATR] variants are visibly higher than thee [-ATR] variants. While the F2 values clearly position the [+ATR] set closer to the peripherals of the vowel space.

6.6.2.1. Is there a quantifiable difference between the advanced and retracted mid-variants?

Hypothesis:

- The null hypothesis for the vowel types/ sample states, there is no difference between the means of the formant resonant values of each vowel type. The average F1 & F2 resonant values are therefore not significantly different between the [+ATR] & [-ATR] variants.

- The null hypothesis for the interaction between the vowel type and thee formant values states, there is no interaction effect between the formant values and the advanced or retracted form. This means that the F1 & F2 token values are consistent across the advanced and retracted variants.

(12) :

Mid-front variants	average		P-value	
	F1	F2	Sample	0,485
ɛ	512,667	2485,233	Interaction	0,000
e	383,3	2653,8		

Mid-back variants	average		P-value	
	F1	F2	Sample	0,015
ɔ	472,909	811,727	Interaction	0,761
o	399,909	718,727		

The statistical comparisons of the mid-advanced and retracted tokens presented in the tableau above indicate different levels of significance. In the vowel charts presented in figure (11) the different ATR mid-variants are positioned distinctly separate within the vowel space. The alpha tokens representing the mean resonant values are also positioned separately in the case of both the front and back vowels. The ellipses inserted in the chart on the right of figure (11) indicates no overlapping of vowel tokens in relation to the front variants, however the back vowel tokens display some visible overlapping. Hence, the alpha tokens appear separate in both the front and back vowel sets, however the back mid-variants display some interlinking between tokens. In the ANOVA calculations for the front mid-variants above, a significant difference is discovered between the variance of the tokens within the sample data sets. Although there is a visible difference between the alpha tokens of these mid-variants sets, the sample p-value in the first tableau is considerably higher than the level of confidence. Hence, the difference between the F1 & F2 mean values of these variants is not significant, which means the initial null hypothesis remains true. However, the disparity between the tokens within the sample data sets of the front mid-variants is both visible within the charts, and calculated as significant according to the F-tests. The interaction p-value in the first tableau is at zero and therefore indicates a high level of significance in the difference between the F1 & F2 token values of the different ATR mid-variants. The front mid-variants therefore do not share a significant difference between their mean formant values, however the formant values of the tokens within each sample data set are not consistent or interchangeable. This signifies a dependency relation between the ATR value of the mid-variants and the formant values of the tokens it comprises.

The back mid-variants then indicate the opposite kind of significance. The first null hypothesis is rejected, as the sample p-values indicates a significant difference between the mean F1 & F2 values of the back mid-variants. The sample p-values is smaller than the level of confidence, and the back ATR variants are therefore considered distinct based on the calculated difference between their total F1 & F2 resonant values. This does not fully correspond with the positioning of the vowels in figure (11) above, as the proximity of the alpha tokens is even closer than the front mid-variants, and these tokens are representative of the mean formant values of each set. However, the ANOVA calculations signify a significant difference between these variants based on the substantial difference between their average formant values. The interaction p-value in the second tableau does not however indicate a significant difference. As stated above, the back mid-variant sets displayed in figure (11) above indicate a visible amount of overlap between tokens, which is also apparent by the interlinked ellipses. The F1 & F2 values of the tokens therefore display a level of consistency across the ATR sets. Hence, the interaction p-value is greater than the level of confidence and supports the second statistical null hypothesis. The back mid-variants share a significant difference between their average formant values, however the F1 & F2 values of their tokens are ultimately consistent.

6.6.3. How is harmony affected by directionality?

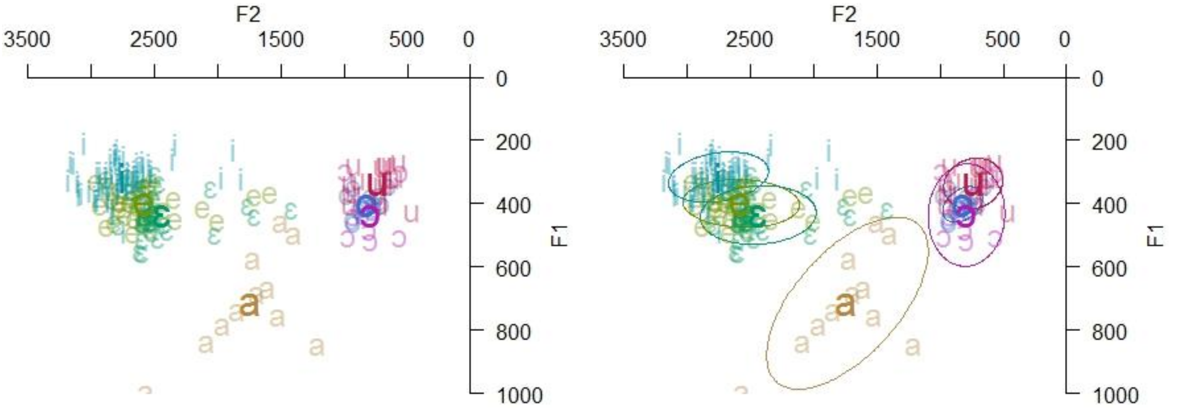
Plotted in the charts below are the formant resonant values of both advanced and retracted mid-vowel sets. However, these segments are exclusively extracted from combinations in which a high vowel is both preceded and followed by a mid-vowel.

(13) Advanced & retracted mid-variants:

a. esilondeni ‘hard covering crust e.g. of bread’
b. esongiweyo ‘savings’

The mid-variants which occur preceding the high vowel are annotated as [+ATR], while those following the high vowel are annotated as retracted. This is aligned with the generalisation which states that [+ATR] spreading occurs regressively and only in the presence of a trigger. Hence, the mid-vowels preceding the high vowel are representative of the acoustic signal of a [+ATR] variant, and those following the high vowel represent the acoustic signal of a [-ATR] variant.

Plotting vowel combination [e/o-i/u-ɛ/ɔ]:



(14) Figure: Plotted in the vowel charts above one may find both the retracted as well as the advanced mid-variant form. These include the results of vowel combinations with a mid-variant both preceding and following a high vowel.
[i] = [+ATR] turquoise; [u] = [+ATR] dark pink; [e] = [+ATR] green; [o] = [+ATR] dark blue; [ɛ] = [-ATR] dark green; [ɔ] = [-ATR] purple; [a] = [-ATR] orange.

In Figure (14) above, two sets of mid-variants are plotted, this includes both the advanced mid-variants [e/o] as well as the retracted forms [ɛ/ɔ]. The measurements plotted here reflect the alternations of the mid-variants in constructions [e/o-i/u-ɛ/ɔ] with a mid-vowel both preceding and following a high vowel. These constructions primarily test whether the generalisation of regressive spreading and leftward directionality are accurate. The distinction between mid-vowel forms is clearly visible in the charts above in which the values for two sets of mid-variant tokens are displayed. Focussing on the front vowels as the visible distinctions appear much clearer, there is an advanced as well as a retracted variant. Considering only the front vowels it seems as though the mid-advanced variant [e] has been inserted neatly between the high vowel [i] and the retracted mid-variant [ɛ]. Therefore, spreading is occurring significantly in one direction, with all mid-variants preceding a high vowel appearing in the advanced form and all those following the high vowel occurring in the retracted form. The same salient distinctions cannot be made for the back mid-variants. The advanced and retracted variants occur in very similar positions within the vowel space, the retracted only occurring slightly lower. However, the retracted variants in this case are not articulated with the tongue in a position as low or toward the back of the mouth as is observed with the default mid-variant in figure (7). What can be seen here is possibly the raising of all mid vowels adjacent to a high vowel, but too a much lesser degree when there is a preceding high vowel.

6.6.3.1. Is there a quantifiable difference between the advanced and retracted mid-variants occurring before and after a high vowel?

Hypothesis:

- The null hypothesis for the vowel types/ sample states, there is no difference between the means of the formant resonant values of each vowel type. The average F1 & F2 resonant values are therefore not significantly different.

- The null hypothesis for the interaction between the vowel type and thee formant values states, there is no interaction effect between the formant values and the advanced or retracted form. This means that the F1 & F2 tokens values are consistent across the advanced and retracted variants.

(15) :

Mid-front variants	average		P-value	
	F1	F2	Sample	0,171
ɛ	435	2429,941	Interaction	0,024
e	401,235	2567,147		

Mid-back variants	average		P-value	
	F1	F2	Sample	0,35
ɔ	435,75	791,125	Interaction	0,737
o	387,5	768,25		

The results of the two way ANOVA statistical analysis presented above indicate both significant as well as insignificant results concerning the difference between the advanced and retracted mid-vowel forms. The first null hypothesis remains true, as the sample p-value is greater than the level of confidence 0, 05. There is therefore no significant statistical difference between the mean formant values of the advanced vowel tokens and those of the retracted vowel tokens. Although a visible difference in F1 & F2 averages may be observed in the tableau above, these value differences are not significant according to the sample F-test. The second null hypothesis then remains true in relation to the back mid-variants, as the interaction p-value is greater than the level of confidence. However, the second null hypothesis is rejected in the case of the front mid-variants, as the interaction p-value is distinctly smaller than the level of confidence. What this means is, there is an interaction effect between the means of the F1 & F2 values and the retracted and advanced mid-front variants. Therefore, the 1st & 2nd formant resonant values of each vowel token are dependent on whether the vowel is advanced or retracted. Hence, as observed in the vowel charts in figure (14) there is a greater discrepancy between the front mid-variant tokens in comparison to the back mid-variant set. This is reiterated in the statistical results, proving a co-dependent relationship between the type of front mid-variant and the formant values.

6.6.3.2. Is there a quantifiable difference between the default mid-variant and the retracted mid-variant?

The resonant values compared in the tableaux below include two different instances of the retracted mid-variant. The variant labelled the default includes mid-vowels extracted from a phonological context in which they are unaffected by any high or low vowel influence. The variants labelled retracted include the retracted variant presented in figure (7) above. This group of mid-vowels are exclusively extracted from the phonological position following a mid-vowel.

(16) Default and retracted mid-variant:

a. ndone 'I did wrong'

b. esil<u>e</u> 'donkey'

The comparison between these variants is necessary in determining whether the retractedness of a segment is dependent on the phonological context. Hence, testing the hypothesis stated above, the retracted variants adjacent to the high vowel may simply be raised to a lesser degree. A statistical comparison of the formant resonant values of these variants will indicate whether there is a significant difference between the retractedness of the mid-variants.

Hypothesis:

- The null hypothesis states, there is no difference between the formant mean values of the default mid-variant and the retracted mid-variant. Both retracted variants share the same average F1 & F2 values.
- The null hypothesis states, there is no interaction effect between the formant values of the vowel tokens and the type of vowel it relates too. Hence, the values of the vowel tokens are consistent across both retracted variants.

(17) :

Mid-variants	average	P-value

	F1	F2	Sample	0,231
Default ɛ	532,529	2424,529	Interaction	0,599
Retracted ɛ	427,235	2383,353		

Mid-variants	average		P-value	
	F1	F2	Sample	0,174
Default ɔ	498,875	855,75	Interaction	0,987
Retracted ɔ	435,75	791,125		

Based on the position of the mid-variants plotted in figure (7) & (14), one may observe a difference between what are both considered retracted mid-variant sets. However, the default variant tokens in figure (7) appear slightly lower and closer to the centre than the retracted set which occur following a high vowel in figure (14). A difference of this nature may be expected, as one might anticipate some co-articulatory effect on the retracted mid-vowel which occurs following the high vowel. A phonetic effect which would not be applicable to a defaulting mid-vowel which occurs without any high vowel influence. In the statistical analysis however this discrepancy is proved to be insignificant. Neither of the null hypothesis can be rejected, they are proven true in the case of both front and back mid-variant sets. The sample p-value in both cases is greater than the level of confidence and therefore preserves the null hypothesis. Hence, there is no significant difference between the mean formant values of the default mid-variant and the retracted mid-variant. Both variants share the same average formant values and are therefore retracted to the same extent. The second null hypothesis also remains true as the interaction p-value in both the default set and the retracted set is greater than the level of confidence. Therefore, there is no dependency relation between the type of retracted variant and the formant values of the token sets. Hence, the formant values measured for the two groups of retracted mid-variant sets are basically interchangeable, whether extracted from a defaulting context or retracted in close proximity to a high vowel. Although there may be a possible difference visible in the vowel charts, this difference is not statistically significant. This result further demonstrates the limitation of any progressive influence of the high vowel trigger.

6.6.4. Is there a significant phonetic difference between the high vowel and the raised mid-vowel?

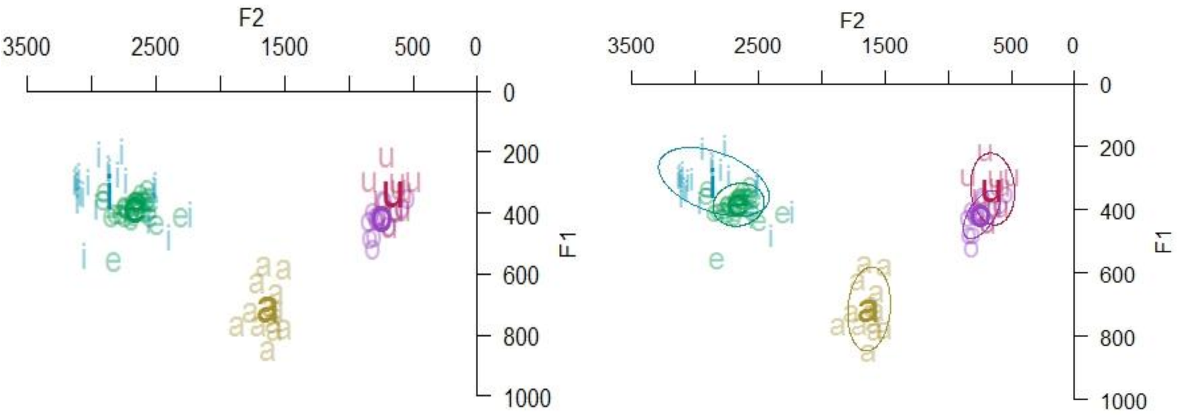
Only one mid-variant is presented in the charts below, which has been extracted from combinations in which the mid-vowel occurs preceding a high vowel. The mid-variant tokens plotted below are therefore indicative of the phonetic effects of raising as a result of the following high vowel trigger.

(18) Advanced mid-variants:

a. ibhedengu 'liar, thief'
b. ixelegu 'dirty person'

The interest in the charts below is between the positioning of the raised mid-variant in relation to the high vowel. The comparison conducted below is necessary in determining whether there is a significant acoustic difference between the [+ATR] mid-variant and the high vowel.

Plotting vowel combination [e/o-i/u]:



(19) Figure: Plotted in the vowel charts above are the mid-variants affected by high vowel raising. These reflect the results from the vowel combinations investigated which include only the measurements of mid-variants preceding a high vowel.
[i] = [+ATR] turquoise; [u] = [+ATR] dark pink; [e] = [+ATR] dark green; [o] = [+ATR] purple; [a] = [-ATR] orange.

Moving on to figure (19) above, these charts display tokens of the advanced mid-variant forms as a result of high vowel raising. The mid-variant data points are therefore measurements extracted from the constructions [e/o-i/u], which include only mid-vowels preceding a high vowel, hence excluding mid-vowels in all other contexts. Immediately there is a significant shift in the position of the mid-variant. Both the front and back mid-variants are situated a lot closer to the high vowel than any retracted mid-variants plotted in the figures above. In terms of their height as well as their F2 resonant value, the mid-variants above are clearly raised due to the presence of the following high vowel. The vowel spaces representative of each sound here are clearly a lot less distinct and significantly interlinked in the case of the front vowels as well as the back vowels. Therefore, the disparity between the high vowels and the mid-vowels is not as prominent in these cases in which the mid-vowels alternate to the advanced form.

6.6.4.1. Statistical comparison between +ATR mid-variants and high vowels

Hypothesis:

- The null hypothesis states, there is no difference between the means of the formant values for the high vowels and those of the advanced mid-vowels
- The null hypothesis states, there is no interaction effect between the formant value means and the [+ATR] variant to which it relates.

(20) :

+ATR Front vowels	average		P-value	
	F1	F2	Sample	0,239
i	311,048	2847,381	Interaction	0,001
e	386,476	2692,429		

+ATR Back vowels	average		P-value	
	F1	F2	Sample	0,009
u	333,7	647,5	Interaction	0,801
o	396,1	722,5		

As the advanced mid-variants and high vowels in figure (19) above are positioned relatively close together, with ellipses which display extensive interlinking, these [+ATR] vowels must be defined as significantly distinct. Consequently, if the advanced mid-variant is not found to be significantly distinct from the high vowel it may be that the [+ATR] variants are true high vowels, rather than alternate mid-variants. As the results indicate different types and levels of significance between the front and back [+ATR] sets, they will be interpreted separately. In the case of the front [+ATR] high and mid-vowel, the first null hypothesis remains true. The sample p-value is greater than the level of confidence and therefore there is no significant difference between the mean values of the formants for the high and mid-front vowels. This result is unsurprising however as the ellipses chart in figure (19) above displays majority of the tokens for the mid-advanced vowel, as well as the alpha token occurring well within the space of the high vowel. Although the formant mean values may be consistent between each vowel, this is not indicative of a completely insignificant difference. The second null hypothesis is definitely rejected, as the interaction p-value is significantly smaller than the level of confidence. Based on the p-value 0,001 there is a highly co-dependent relationship between the formant values of each group of tokens and the particular vowel they represent. Hence, the formant means for each [+ATR] front vowel may not indicate a significant difference, however there is a clear interaction between the formant values and the vowel to which they relate.

The ANOVA calculations for back [+ATR] vowels display a different type of significance to that of the front [+ATR] vowels. The first null hypothesis is confidently rejected, as the sample p-value is smaller than the level of confidence. Therefore, there is a significant difference between the formant value means of the [+ATR] high back vowel and the [+ATR] mid-back vowel, which indicates that the raised mid-vowels are not acoustically the same as the high vowels. On the other hand the interaction p-value is greater than the level of confidence and therefore the second null hypothesis remains true. This means there is no significant interaction effect between the formant mean values and the vowel to which they relate. Again this discrepancy may be observed in figure (19), unlike the front mid-variant, the back mid-variant tokens appear to occur less intertwined in the space of the high back vowel. Therefore, it is expected that their formant mean values would be significantly different. It is however clear in each case that there is some significant discrepancy between the high vowels and the advanced mid-vowels. Based on the statistical analysis above the mean value of the vowel tokens may be inconsistent and thereby display a significant discrepancy. This is clear in the comparison between the back vowels which do not share the same F1 & F2 averages. As well as the relation between the formant values of the tokens and the vowel type which demonstrate a level of co-dependence. As in the case of the front vowels, which share the same F1 & F2 averages but indicate a significant relationship between the individual token values and the vowel type.

6.6.5. Does the low vowel [a] block spreading?

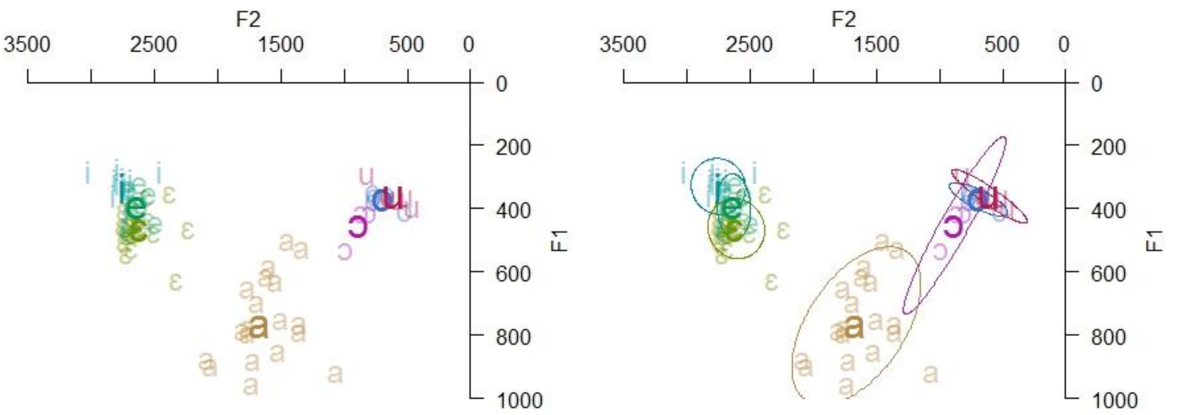
Plotted in the charts below are two sets of mid-variants. The [-ATR] mid-variants are of particular interest as they have been blocked from raising and are therefore retracted. Hence, due to the phonological position of the low vowel the retracted mid-variants below could not undergo vowel raising. Extracted from the vowel combination [ɛ/ɔ-a-i/u], the [+ATR] mid-variants simply include those which did not occur preceding the low vowel and were triggered for raising.

(21) Advanced & retracted mid-variants:

a. ebeqabukelisa ‘to look back’
b. esandleni ‘applause’

The mapping of these variants in the charts below is necessary to measure the extent of the retraction of blocked mid-vowels. Hence, are the retracted mid-variants which occur as a result of blocking, further retracted due to co-articulatory effects.

Plotting vowel combination [ɛ/ɔ-a-i/u]:



(22) Figure: The vowel charts above include the retracted mid-variant which occurs as a result of blocking by the low vowel [a]. The significance of this chart is the position of the retracted variant in comparison to those extracted from the vowel combinations reflected in the charts above.
[i] = [+ATR] turquoise; [u] = [+ATR] dark pink; [e] = [+ATR] dark green; [o] = [+ATR] dark blue; [ɛ] = [-ATR] green; [ɔ] = [-ATR] purple; [a] = [-ATR] orange.

Presented in the chart above are both the mid-advanced and mid-retracted variants as displayed in figure (14). However, the retracted mid-variants plotted in figure (22) above have shifted even lower relative to those presented in figure (14). When spreading is blocked by the low vowel [a] the preceding mid-vowel is expected to surface as the retracted variant, which is demonstrated by the values for segments [ɛ/ɔ] in the charts above. The retracted variants displayed in these charts are limited to those which reflect cases of blocking. When compared to the retracted mid-variants following a high vowel in the chart in figure (14), there is a visible difference in the height of the vowels now preceding the low vowel [a], and those which are not affected by the occurrence of a low vowel. The vowel combinations from which the mid-variants above were extracted include only [ɛ/ɔ-a-i/u]. Although the mid-variants preceding the high vowel trigger should always surface as the mid-advanced variant, those first preceding the low vowel and then the high vowel are expected to occur in the retracted form. What appears to be happening is perhaps co-articulation between the mid-vowels which occur adjacent to either a high or low vowel. This deduction is based on the assimilation of the mid-variants in the vowel charts above to the nearest adjacent high or low vowel. Again the ellipses display interlinking between the vowel spaces of the mid-variants, specifically between the advanced mid-variants and the high vowel.

6.6.5.1. Statistical difference between the mid-vowels being blocked and those considered to be the default

Hypothesis:

- The null hypothesis states, there is no difference between the mean formant values of the default mid-variant and the retracted mid-variant which has undergone blocking.

- The null hypothesis states, there is no interaction effect between the mean values of the formants and the vowels to which they relate.

(23) :

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,247
Default ɛ	532,529	2424,529	Interaction	0,003
Blocked ɛ	451,471	2605,294		

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,115
Default ɔ	504,167	862,5	Interaction	0,977
Blocked ɔ	417,667	779		

The statistical comparison of the default mid-variant and the retracted mid-variant as a result of blocking displayed above, is indicative of a partially significant discrepancy. The initial null hypothesis remains true for both the front default and blocked mid-variants, as well as those articulated in the back of vowel space. Both tableaux display a sample p-value greater than the level of confidence, hence there is no significant difference between the formant mean values of the default mid-variant and the blocked mid-variant. There is however a significant interaction effect between the F1 & F2 mean values and the defaulting and retracted mid-variants. Therefore, although there is not a significant discrepancy between the mean values of the formants, there is co-dependence between these variables and the vowel to which they relate. In the case of the back mid-variants however, the null hypothesis remains true. There is no significant interactive relation between the formant mean values and the related vowel. Considering the possible effects of co-articulation, one would expect the retracted mid-variant as a result of blocking, to occur lower and closer to the centre of the vowel space. Therefore, due to its close proximity to the low vowel [a] the retracted mid-variant should occupy a position closer to the influencing segment. However, the average F1 values indicate that the blocked mid-variants would occupy a slightly higher position within the vowel space due to their lower F1 value. A comparison of the average F2 values indicate that the default mid-variants are slightly more centred than the blocked mid-variants. Therefore, there appears to be little or no effects of co-articulation with the adjacent low vowel in the case of retracted mid-variants due to blocking.

6.6.6. Is the retracted form preceding [a] lower than the default form?

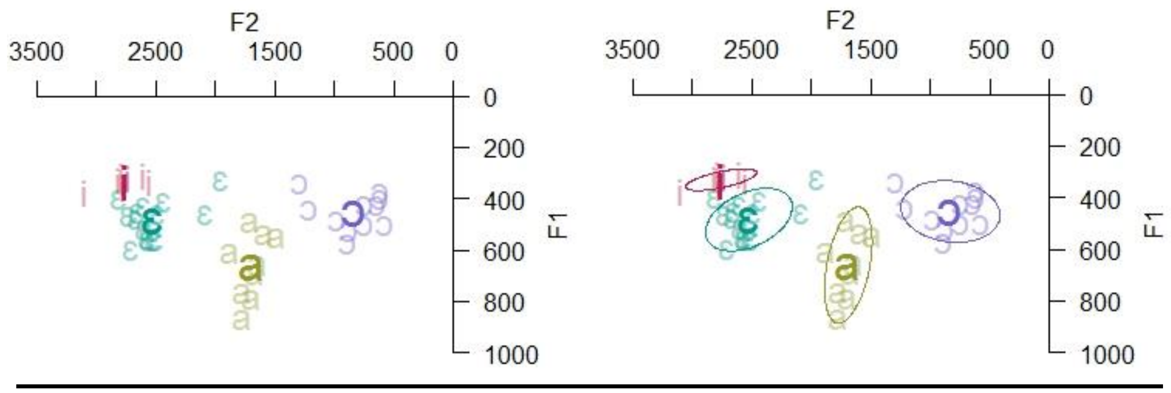
The retracted tokens plotted in the charts below are not the same as those preceding the low vowel in figure (22) above. The retracted mid-variants plotted below are extracted from the vowel combination [ɛ/ɔ-a]. Retracted as a result of defaulting and not due to low vowel blocking. Hence, in this combination the low vowel is not followed by a high vowel and therefore the mid-variant was not intended to be raised.

(24) Retracted mid-variants:

a. feketha ‘play/trifle’
b. godola ‘feel cold’

The following mapping is again used to test the co-articulatory effect of the low vowel on a preceding mid-vowel. Although the mid-variant extracted from these combinations should occur in the default [-ATR] form, are they retracted to the same extent as the default form presented in figure (7) above.

Plotting vowel combination [ɛ/ɔ-a]



(25) Figure: The vowel charts above include the retracted mid-variant which occur preceding the low vowel [a]. The significance of this chart is the position of the retracted variant in comparison to those extracted from the vowel combinations reflected in the charts above.
[i] = [+ATR] dark pink; [ɛ] = [-ATR] turquoise; [ɔ] = [-ATR] purple; [a] = [-ATR] yellow.

The measurements plotted in the vowel charts above include regular occurrences of the high and low vowels, while the mid-variants include only those which occur preceding the low vowel [a]. This does not include cases of blocking, but simply cases in which the mid-vowel has occurred in close proximity to the low vowel. Thereby testing for an effect of /a/ on /e/, independent of other harmony. The mid-variant should therefore already appear in the retracted form by default, as raising has not occurred. However, to test the effects of the low vowel as a possible harmonic trigger or articulatory influence all preceding mid-variants have been measure and plotted above. The formant values were extracted from the constructions with the vowel combination [ɛ/ɔ-a]. The mid-vowels recorded in these constructions are expected to occur in the retracted form, however as a consequence of the following low vowel they are expected to occur slightly lower and more centered in comparison to the default forms plotted in figure (7). However, no articulatory effects of this nature are evident based on the vowel charts above.

6.6.6.1. Is there a significant difference between the default mid-variant and the retracted mid-variant preceding a low vowel?

Hypothesis:

- The null hypothesis states, there is no difference between the formant mean values of the default mid-variants and the retracted mid-variants preceding a low vowel.

- The null hypothesis states, there is no interaction effect between the formant mean values and the vowel to which they relate.

(26) :

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,49
Default ɛ	537	2404,625	Interaction	0,087
ɛ-a	485	2525,563		

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,972
Default ɔ	472,909	811,727	Interaction	0,663
ɔ-a	453,545	834,545		

Based on expectations of co-articulation and a slight visible difference in the positioning of mid-vowels plotted in figure (7) and figure (25) above, one would expect discrepancies in the F1 & F2 values of these mid-variants. As a consequence of the effects of the low vowel, one would expect the preceding retracted mid-variant to occur slightly lower and more centered within the vowel space. Hence, there would be a closer assimilation to the position of the following low vowel, which would not be expected in the case of the default mid-variant. This co-articulatory hypothesis is however disproved by the statistical analysis above, as neither of the null hypothesis may be rejected. The sample p-value in both tableaux is greater than the level of confidence and therefore supports the initial null hypothesis, proving no significant difference between the mean formant values of the default variant in comparison to the retracted variant preceding the low vowel. The interaction p-values, also greater than the level of confidence reiterates that there is no dependency relation between the formant resonant values observed in each set, and the particular vowel to which they relate. Furthermore, when compared, the average F1 & F2 values differ only slightly between the front and back mid-variants. Hence, the expected co-articulatory effects are not significant based on the statistical analysis.

6.6.7. Is the raised mid-vowel affected by a low vowel preceding the high vowel?

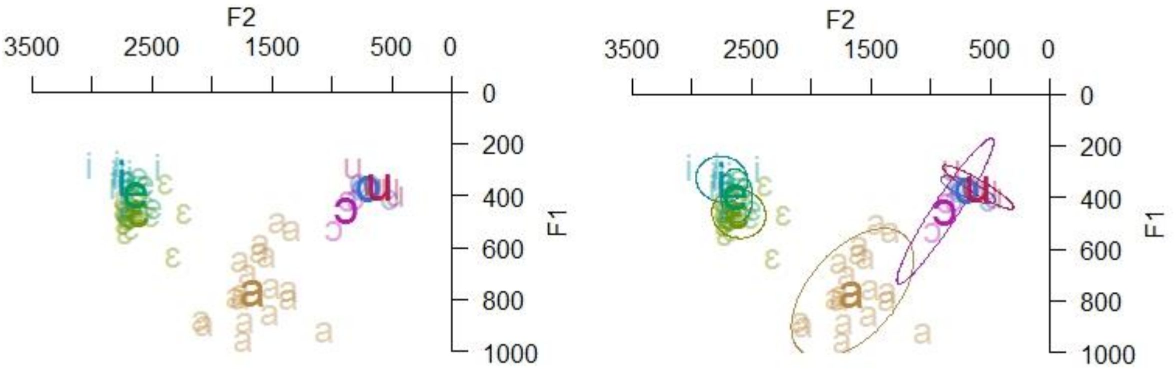
Plotted below are the mid-variants extracted from the combination [e/o-i/u-a]. This includes instances of mid-vowels preceding a high vowel trigger which is then also preceded by a low vowel. This combination is the reverse of that used to determine the effects of blocking, as the high vowel now precedes the low vowel and can therefore spread [+ATR] across all preceding mid-vowels. The pertinent mid-variants below are therefore the [+ATR] forms which have assimilated for [+ATR] due to the following high vowel trigger.

(27) Advanced mid-variants:

a. ebebheculula 'he/she opened the eyelid with the finger'
b. ebephengulula 'he/she turn upside down'

The significance of mapping the advanced mid-variants extracted from this particular combination is to determine whether the extent of raising is at all affected by the proximity of a following low vowel. The comparisons above test the effects of the low and high vowels on the retractedness of mid-variants extracted from different phonological contexts. Here the extent of raising is measured in the context of a high vowel trigger further followed by a low vowel.

Plotting vowel combination [e/o-i/u-a]



(28) Figure: The vowel charts above include the advanced mid-variants which occur preceding a high vowel trigger. These advanced mid-variants are still in close proximity to the low vowel [a], as the high vowel trigger is preceded by the low vowel. The significance of this chart is the position of the retracted variant in comparison to those extracted from the vowel combinations reflected in the charts above.
[i] = [+ATR] turquoise; [u] = [+ATR] dark pink; [e] = [+ATR] dark green; [o] = [+ATR] dark blue; [ɛ] = [-ATR] green; [ɔ] = [-ATR] purple; [a] = [-ATR] orange.

The vowel charts above are of interest as they display cases of mid-vowel forms which are followed by both a high vowel trigger, as well as a low vowel. The mid-variants undergo raising as a result of the high vowel trigger, which is then directly followed by a low vowel [a]. Therefore, although these mid-variants are expected to appear in the advanced form, it is possible that the close proximity of the low vowel may affect the formant measurements of the mid-variant, altering their position within the vowel space. Extracted from constructions with the vowel combination [e/o-i/u-a], the raised mid-variants [e,o] are therefore the relevant mid-variant forms plotted in the chart above. In comparison to the advanced mid-variants plotted in figure (14) above, these tokens appear to occur visibly closer to the high vowels. Hence, the results displayed in the vowel charts demonstrate no consequent effects of proximity to the low vowel. The mid-vowels appear to be raised by the high vowel trigger and are then generally unaffected by any further phonetic influence.

6.6.7.1. Compare the extent of raising in the phonological context of a high vowel trigger followed by a low vowel

Hypothesis:

- The null hypothesis states, there is no difference between the mean formant values of the raised vowels & the raised vowels in close proximity to a low vowel.

- The null hypothesis states, there is an interaction effect between the mean formant values and the particular vowel to which they relate.

(29) :

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,071
Raised e	386,476	2692,429	Interaction	0,035
e-i/u-a	400,714	2509,810		

Mid-vowel variants	average		P-value	
	F1	F2	Sample	0,070
Raised o	413,75	782,5	Interaction	0,052
o-i/u-a	407,25	942,25		

The results of the statistical analysis above clearly show no significant difference between the mean formant values of the raised vowel in comparison to the raised vowel in close proximity to a low vowel. With both sample p-values greater than the level of confidence, the initial null hypothesis is generally proven true. As displayed in the tableau above there are only slight discrepancies in the average F1 & F2 values of the front and back mid-variants, hence they are generally consistent across each of these mid-variant forms. The second null hypothesis may however be rejected in the case of the front mid-variants as the interaction p-value is smaller than the level of confidence. This therefore indicates that although the mean formant values may be consistent between the front mid-variants, there is a significant dependency relation between the formant values and the relevant vowel. Worth noting, the interaction p-value relating to the back vowel is only just greater than the level of confidence. Therefore, it may be likely with a larger data set that a significant interaction may be identified. However, the ANOVA results above indicate only a significant discrepancy between the front mid-variants.

6.7. Discussion

As predicted the mid-vowels which are annotated as advanced or retracted, are situated as such within the vowel space. The recorded mid-advanced variants prove to have lower frequency values for the first formant in comparison to the retracted mid-variants, and therefore appear higher on the vowel chart, this is illustrated in figure (11), (14) & (19) above. The insertion of ellipses into each of the vowel charts generally serve as a calculated indication of distributive similarities between vowels. This calculation is based on one standard deviation of the estimated data and the co-variance of the tokens within each sample set (McCloy 2012). Specifically, a significant overlap is observed in the distributive location of the mid- advanced variants and the high vowels. However, as demonstrated in the statistics tableau (20) there is a significant difference between the formant values of the high vowels and advanced mid-vowels, as well as a significant level of interaction between the formant values of the tokens in each data set and the vowel to which they relate. Therefore, the visible distributive overlaps between the high vowels and the advanced mid-variants may be indicative of acoustic similarity, however the statistical analysis demonstrates a significant difference between these vowels. Hence, the mid-advanced variants are distinct segments acoustically idiosyncratic and separate from the high [+ATR] segments.

The disparity between the high vowels and the mid-retracted variants is quite significant and shows no interlinking of vowel spaces. This further iterates a significant difference between the [+ATR] & [-ATR] mid-variants. As displayed in the array of vowel charts above, there are phonetic situations in which mid-vowels may be influenced by surrounding segments. Based on the statistical results in tableaux (23) and (29) above, there appears to be a stronger co-articulatory effect resulting from the proximity of a high vowel as in the combination [ɛ/ɔ-a-i/u]. The retracted vowels in this case have not been proven statistically different from the default mid-variant, however the average formant values indicate a slightly higher and less centred position. The opposite combination however, testing the co-articulatory effects of the low vowel [e/o-i/u-a], indicate no change in acoustic composition. The raised vowel in close proximity to the low vowel does not occur lower or more centred within the vowel space to assimilate to the position of the low vowel, both the F1 & F2 average values are completely consistent. Therefore, it appears the proximity of the high vowel produces a stronger articulatory effect than that of the low vowel. Overall the default mid-variants which occur in the combination [ɛ/ɔ] uninfluenced by any preceding or following segments, are positioned lower and more centred than any other recorded mid-variant. Therefore, the mid-variants most isolated from the high vowel within the vowel space are not the systematically retracted or blocked mid-variants, but the mid-variants which are secluded within the prosodic word, see figure (7).

The low vowels are scattered at lower points on the graph because of their high F1 values, while the higher vowels all appear within the upper level of the graphs which include the mid-advanced variants. The mid-variants are visibly clustering in specific and limited vowel spaces depending on whether they have been annotated as advanced or retracted. Many of the generalisations presented by Jokweni & Thipa (1996) are strengthened by this analysis, simply because the predicted annotation resulted in data which revealed significant patterns. Hence, claims about vowel raising and low vowel blocking are discovered to be accurate in that there is a visible phonetic difference in the acoustic composition of these separate mid-variant sets. However, based on further analysis of these findings, it appears that harmony may be working along a spectrum. The vowels annotated as retracted are not always retracted to the same extent. Those which occur preceding the low vowel, surprisingly appear to have a slightly lower F1 value than those considered the default option, uninfluenced by either high or low vowels. This suggests that the process these mid-vowels undergo may not be strict harmony as presented in previous literature, but may include a force of phonetic co-articulation. Therefore, the vowels situated at the edges of the vowel space will inherently impact the acoustic nature of the following or preceding mid-vowels, as they are more flexible in the intermediate spaces. It is apparent however that the extent of this co-articulation may vary, and this is why we find that the acoustic data appears to support the generalisations presented by Jokweni & Thipa (1996)

7. Conclusion

The vowel harmony system in isiXhosa shares many similarities with those observed in other southern Bantu, Nilo-Saharan and Nilo-Congo languages (Casali 2008). Therefore, as reported in Jokweni and Thipa (1996) the two sets of mid-vowels found in isiXhosa are distinguished for the feature ATR. This was confirmed by the correlation observed between the formant resonant frequencies and the size of the pharyngeal cavity. The overall results of the acoustic analysis presented above reveal all [+ATR] vowels possess a lower F1 value, while all [-ATR] vowels possess a higher F1 value. According to previous acoustic investigations of ATR the lower the F1 value the bigger the pharyngeal cavity and vice versa (Halle & Stevens 1969). Hence, the ATR contrast is considered an alternation in the shape and size of the pharyngeal cavity. Although the F2 resonant values have been observed to vary systematically cross linguistically, the result found in isiXhosa is consistent with that found in Nilo-Congo languages (Fulop 1998). The overall results of the F2 resonant values indicate that the [+ATR] vowels are positioned closer to the peripheral of the vowel space than the [-ATR] vowels. This form of acoustic analysis remains the most established and reliable indication of ATR contrast (Guion 2004). Evidence of possible co-articulation is limited to a non-significant observation of [-ATR] mid-variant assimilation to proximal high vowels. Overall the results confirm a phenomenon with a stronger resemblance to harmony rather than co-articulation as proposed in SiSwati (Malambe 2015). All further harmony specifications presented in Jokweni & Thipa such as raising, blocking, directionality, and the iterative nature of spreading are supported by the acoustic results.

The harmony process has furthermore been successfully analysed and characterised using Optimality theory. The original no-disagreement approach to OT was initially applied to the harmony system. This approach could not however account for all of the complexities of this particular system, hence an adaptation to the constraint set was necessary. The problematical idiosyncrasies of this system include a combination of requirements; [+ATR] spreading may only be triggered by a high vowel, a novel finding which has not been reported in previous literature; blocking caused by the low vowel [a] must be accounted for by [+ATR] progressive spreading; while the correct directionality of [+ATR] spreading is accounted for by enforcing regressive spreading (Sasa 2009). The combination of these tasks have proven to require a more nuanced approach to OT, one in which the constraints driving harmony are further restricted. Hence, the issue is that harmony must be forced both progressively as well as regressively in order to account for the complexities of the pattern (Pulleyblank 2002). However, this results in a contradictory ranking, one in which the attested form cannot be derived in every instance. No ranking of these harmony constraints can therefore result in the successful derivation of the harmony pattern. The adaptation to the no-disagreement approach therefore encompasses the addition of a second feature into the basic spread constraint. This is aligned with the concept of correspondence theory, in which only segments retaining a particular feature are forced to harmonize (Krämer 2001). As the low vowel is not involved in the raising process of this harmony system, the spread constraint is restructured and bound by the feature [-Low]. Therefore, instances in which both raising as well as blocking occur are accounted for by the ranking, as well as those instances in which only raising occurs and no progressive spreading is required. This adaptation to the harmony constraints ensures that spreading will not be enforced both progressively and regressively. The regressive spread constraint may now only be activated by the occurrence of an appropriate trigger, specifically one retaining the feature [-Low].

The harmony system reported to exist in isiXhosa is proven to be more complex than those recorded in other African languages. Based on the resonant frequency values analysed and compared across the different vowel segments, the relevant harmonic feature is advanced tongue root (ATR). The feature [+ATR] is considered dominant as it actively spreads within a defined phonological domain, this is similar to many languages which share an ATR contrast (Casali 2008). Segments are therefore altered to [+ATR] at the expense of [-ATR] forms. Although [+ATR] is reported to be dominant in numerous ATR harmony systems. Those considered to be four height systems in which only mid-vowels alternate, generally retain a system in which [-ATR] serves as the dominant value (Bakovic 2000, Guion 2004). Hence, the harmony system described above share both similarities and differences in comparison to other ATR harmony systems.

Questions left open for further research, the issue of particular consonantal interference discussed in the data & generalisations chapter. Further acoustic investigation would provide a more insightful explanation of the effects of nasal consonants on preceding mid-vowels. To answer the question of whether this influence is entirely consistent, and whether the closure impressed on the preceding mid-vowel is a consequence of the reconfiguration of the vocal tract to produce the nasal. An extended investigation might focus on the effect of the lateral approximant [l] on any following mid-vowels. Again the vowel which should occur in the retracted form is observed to have an acoustic signal which is indicative of a ‘closed’ rather than ‘open’ variant. In both of these instances [+ATR] values are imposed upon segments which are meant to retain a [-ATR] value. Interestingly the reverse effect does not occur.

References

Amir, O., Wolf, M., & Amir, N., 2009. A clinical comparison between two acoustic analysis softwares: MDVP and Praat. Biomedical Signal Processing and Control, 4(3), 202-205.

Archangeli, D. and Pulleyblank, D., 1989. Yoruba vowel harmony. Linguistic inquiry, 20(2), pp.173-217.

Archangeli, D. and Pulleyblank, D., 1992. Grounded Phonology, ms. University of Arizona and University of British Columbia.

Archangeli, D., 1997. Optimality Theory: an introduction to linguistics in the 1990s. In Archangeli, Diana B. & Pulleyblank, Douglas G. (eds.), Optimality Theory: an overview. Malden, MA: Blackwell. 1–32.

Archangeli, D. and Pulleyblank, D., 2002. Kinande vowel harmony: domains, grounded conditions and one-sided alignment. Phonology, 19(02), pp.139-188.

Baayen, R.H., 2008. Analyzing linguistic data: A practical introduction to statistics using R. Cambridge University Press.

Bakovic, E., 2000. Harmony, Dominance and Control. Rutgers University Ph.D. dissertation (Rutgers Optimality Archive #360).

Baković, E., 2007. Local assimilation and constraint interaction. The Cambridge Handbook of Phonology, (1), pp.335–352

Beckman, J.N., 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. Phonology, 14(01), pp.1-46.

Boersma, P. and Weenink, D., 2005. Praat, software for speech analysis and synthesis. Available from: <http://www.praat.org>, 3/2005.

Bluman, A.G., 2000. Elementary statistics. McGraw Hill Publishers.

Chandler, D., 1987. Introduction to modern statistical mechanics. Introduction to Modern Statistical Mechanics, by David Chandler, pp. 288. Foreword by David Chandler. Oxford University Press, Sep 1987. ISBN-10: 0195042778. ISBN-13: 9780195042771, p.288.

Brasoveanu, A. and Prince, A., 2011. Ranking and necessity: The fusional reduction algorithm. Natural Language & Linguistic Theory, 29(1), pp.3-70.

Cole, J. and Kisseberth, C., 1994. An optimal domains theory of harmony. Studies in the Linguistic Sciences, 24(2), pp.101-114.

Clements, G.N., 1985. Akan vowel harmony: a nonlinear analysis. African Linguistics: Essays in Honor of WMK Semikenke, pp.55-98.

Doke, Clement M., 1926. “The Phonetics of the Zulu Language.” Bantu Studies 11: 1–138.

Doke, Clement M., 1954. The Southern Bantu Languages. Handbook of African Languages. New york Capetown: Oxford University Press

Fitzgerald, F.S., 2009. The beautiful and damned. Oxford University Press.

Fowler, C.A., 1980. Coarticulation and theories of extrinsic timing. Journal of Phonetics, 8(1), pp.113-133.

Fulop, S.A., Kari, E. and Ladefoged, P., 1998. An acoustic study of the tongue root contrast in Degema vowels. Phonetica, 55(1-2), pp.80-98.

Gafos, A.I., 2014. The articulatory basis of locality in phonology. Routledge.

Gay, T., 1977. Articulatory movements in VCV sequences. The Journal of the Acoustical Society of America, 62(1), pp.183-193.

Guion, S.G., Post, M.W. and Payne, D.L., 2004. Phonetic correlates of tongue root vowel contrasts in Maa. Journal of Phonetics, 32(4), pp.517-542.
Halle, M., 1962. Phonology in generative grammar. Word, 18(1-3), pp.54-72.
Halle, M. and Stevens, K., 1969. On the feature 'advanced tongue root'. Quarterly Progress Report of the MIT Research Laboratory in Electronics, 94, pp.209-215.
Halle, M., 1995. Feature geometry and feature spreading. Linguistic inquiry, 26(1), pp.1-46.
Harris, J., 1987. Non-structure-preserving rules in Lexical Phonology: Southeastern Bantu harmony. Lingua, 72(4), pp.255-292.
Hyman, L.M., 2013. Enlarging the scope of phonologization. Origins of sound change: Approaches to phonologization, pp.3-28.
Jokweni, M.W. and Thipa, H.M., 1996. ATR harmony in Xhosa. South African Journal of African Languages, 16(4), pp.119-123.
Jordan, A.C., 1966. A practical course in Xhosa 1st ed., Wisconsin: Longmans Southern Africa (PTY) limited.
Jurgec, P., 2011. Feature spreading 2.0: A unified theory of assimilation. University of Tromsø Ph.D. dissertation.
Kenstowicz, M.J., 1994. Phonology in generative grammar (pp. 195-249). Oxford: Blackwell.
Khabanyane, K.E., 1991. The five phonemic vowel heights of Southern Sotho: an acoustic and phonological analysis. Working Papers of the Cornell Phonetics Laboratory, 5, pp.1-36.
Kimper, W.A., 2011. Competing triggers: Transparency and opacity in vowel harmony. University of Massachusetts Amherst Ph.D. dissertation.
Krämer, M., 2001. Vowel harmony and correspondence theory. Heinrich-Heine-Universität Düsseldorf Ph.D. dissertation.
Kurisu, K., 2000. Richness of the base and root fusion in Sino-Japanese. Journal of East Asian Linguistics, 9(2), pp.147-185.
Ladefoged, P., 2003. Phonetic data analysis: An introduction to fieldwork and instrumental techniques. Malden, MA: Blackwell.
Ladefoged, P. and Disner, S.F., 2012. Vowels and consonants. John Wiley & Sons.
Le Roux, M. and Le Roux, J., 2008. An acoustic assessment of Setswana vowels. South African Journal of African Languages, 28(2), pp.156-171.
Liljencrants, J. & Lindblom, B., 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. Language, 48, pp.839-862.
Lindau, M., 1978. Vowel features. Language, 54, pp.541-563.
Maduka, O.N., 1988. Size and shape ideophones in Nembe: A phonosemantic analysis. Studies in African linguistics, 19(1), pp.93-113.
Mahanta, S., 2008. Directionality and locality in vowel harmony: With special reference to vowel harmony in Assamese. Netherlands Graduate School of Linguistics.
Malambe, G.B., 2015. Mid vowel assimilation in siSwati. *Southern African Linguistics and Applied Language Studies*, *33*(3), pp.261-272.
McCarthy, J. and Prince, A., 1995. Faithfulness and reduplicative identity. Papers in optimality theory, ed. by Jill Beckman, Laura Dickey, and Suzanne Urbanczyk, 249–384. University of Massachusetts occasional papers in linguistics, 18.
McCarthy, J.J., 2002. A thematic guide to Optimality Theory. Cambridge University Press.
McCarthy, J.J., 2011. Autosegmental spreading in optimality theory. Tones and features: Phonetic and phonological perspectives, pp.195-222.

McCloy, D., 2012. phonR: R tools for phoneticians and phonologists. R package version 1.0.7. http://drammock.github.io/phonR/
Öhman, S.E., 1966. Coarticulation in VCV utterances: Spectrographic measurements. The Journal of the Acoustical Society of America, 39(1), pp.151-168.
Owren, M. J., 2008. GSU Praat Tools: Scripts for modifying and analyzing sounds using Praat acoustics software. Behavior research methods, 40(3), 822-829.
Padgett, J., 1997. Partial class behavior and nasal place assimilation. In K. Suzuki & D. Elzinga (eds.) Proceedings of the Southwest Optimality Theory Workshop. Tucson: University of Arizona. 1-40.
Pahl, H. W., 1967. IsiXhosa (Xhosa). Johannesburg: Educum Publishers.
Perkins, J., 2005. The RTR harmonic domain in two dialects of Yorùbá. The University of British Columbia Ph.D. dissertation.
Petrova, O., Rosemary, P., Catherine, R. and Szilrd, S., 2000. Voice and Aspiration in OT. Paper presented at the 74th Annual Meeting of the Linguistic Society of America, Chicago.
Poletto, R., 1999. Default lexical tone interaction in Runyankore verbs. New Dimensions in African Linguistics and Languages, 3, p.107.
Prince, A. and Smolensky, P., 2004. Optimality Theory: Constraint Interaction in Generative Grammar. MIT Press, Cambridge, MA.
Pulleyblank, D., 2002. Vowel Height. In Grounded Phonology. Massachusetts: The MIT Press, pp. 491–518.
Pulleyblank, D., 2002, August. Harmony drivers: No disagreement allowed. In Annual Meeting of the Berkeley Linguistics Society(Vol. 28, No. 1, pp. 249-267).
Reetz, H. and Jongman, A., 2009. Phonetics. Transcription, production, acoustics and perception. Wiley Blackwell.
Sapir, E., 1929. A study in phonetic symbolism. Journal of experimental psychology, 12(3), p.225.
Sasa, T., 2006. Directionality in harmony: a case study of Pulaar. Paper presented at the 14th Manchester Phonology Meeting, Manchester, UK.
Sasa, T., 2008. A Span-Theoretic account of Pulaar ATR harmony. Paper presented at the 14th Mid Continental Workshop on Phonology, Minneapolis, MN.
Sasa, T., 2009. Treatments of vowel harmony in Optimality Theory. The University of Iowa.
Schadeberg C., 2003. Derivation. In The Bantu languages, Derek Nurse & Gerard Philippson (eds), London: Routledge, 71-89
Sibanda, G., 2009. Vowel processes in Nguni: Resolving the problem of unacceptable VV sequences. In Selected Proceedings of the 38th Annual Conference on African Linguistics (pp. 38-55). Somerville, MA: Cascadilla Proceedings Project.
Smolensky, P., 1993, October. Harmony, markedness, and phonological activity. In Rutgers Optimality Workshop (Vol. 1, pp. 87-0000).
Starwalt, C.G.A., 2008. The acoustic correlates of ATR harmony in seven-and nine-vowel African languages: A phonetic inquiry into phonological structure. The University of Texas at Arlington Ph.D. dissertation.
Suzuki, K., 1998. A typological investigation of dissimilation. The University of Arizona Ph.D. dissertation.
Tshabe, S.L., Shoba, F.M., Mini, B.M., Pahl, H.W., Pienaar, A.M. and Ndungane, T.A., 2006. The greater dictionary of isiXhosa (Vol. 3). University of Fort Hare.

Van der Hulst, H. and Van de Weijer, J., 1995. Vowel harmony. The handbook of phonological theory, ed. by John A. Goldsmith, 495-534.
Van de Weijer, J., 1996. Segmental structure and complex segments. Tübingen:
Max Niemeyer Verlag (Vol. 350). Walter de Gruyter GmbH & Co KG.
Walker, R., 2009, January. Similarity-sensitive blocking and transparency in Menominee. In 83rd Annual Meeting of the Linguistic Society of America, San Francisco.
Walker, R., 2016, June. Surface correspondence and discrete harmony triggers. In Proceedings of the Annual Meetings on Phonology (Vol. 2).
Wilson, C., 2006, June. Unbounded spreading is myopic. In Current Perspectives on Phonology workshop, Phonology Fest. Indiana University, June, 2006.
Wiswall, W.J., 1991. Partial vowel harmonies as evidence for a Height Node. The University of Arizona Ph.D. dissertation.