

Vowel Harmony in Tuvan and Igbo:  
Statistical and Optimality Theoretic Analyses

December 8, 2010

Swarthmore College

Amy Smolek

## Table of Contents

Acknowledgements.....	4
0. Abstract.....	5
1. Introduction to Vowel Harmony.....	5
1.1 Languages that have VH.....	7
1.2 Reasons for VH.....	9
1.3 VH System Change.....	11
2. Tuvan.....	12
Table 2.1 Tuvan Vowel Distribution.....	12
Table 2.2 Vowel Feature Distribution.....	13
Table 2.3 Plural Suffix Alternation in Tuvan.....	14
2.1 Disharmony in Tuvan.....	15
Table 2.4 Exceptions in Tuvan.....	15
3. Igbo.....	16
Table 3.1 Igbo Vowel Distribution.....	16
3.1 Disharmony in Igbo.....	17
4. Analysis.....	18
4.1 Vowel Harmony Calculator.....	18
Figure 4.1 VHC Screenshot.....	19
5. Tuvan Data.....	20
Table 5.1 Tuvan Vowel Correspondences.....	21
5.1 Tuvan VHC Results.....	21
Table 5.2 Tuvan VHC Results Summary.....	21
Table 5.3 Disharmonic Words in Tuvan.....	22
Table 5.4 Disharmonic Words in Turkish.....	23
6. Igbo Data.....	24
Table 6.1 Igbo Vowel Correspondences.....	25
6.1 Igbo VHC Results.....	25
Table 6.2 Igbo VHC Results Summary.....	25
Table 6.3 Disharmonic Words in Igbo.....	26
Table 6.4 Igbo VHC Results Summary, Minus Disharmonic Suffixes.....	27
Table 6.5 Igbo Print Sources VHC Results Summary.....	28

Table 6.6 Igbo Print Sources VHC Results Summary, Minus Disharmonic Suffixes..... 28

7. Statistical Summary..... 29

8. Optimality Theory and Harmony..... 30

    8.1 Faithfulness Constraints..... 31

    8.2 Markedness Constraints..... 32

9. Tuvan OT Analysis..... 35

    Tableau 9.1..... 35

    Tableau 9.2..... 35

    Tableau 9.3..... 35

    Tableau 9.4..... 35

    Tableau 9.5..... 36

    Tableau 9.6..... 36

10. Igbo OT Analysis..... 37

    Tableau 10.1..... 37

    Tableau 10.2..... 37

    Tableau 10.3..... 38

11. OT Summary..... 38

12. Conclusion..... 38

Bibliography..... 40

Data Sources..... 43

Appendix A: Tuvan Disharmony Log..... 45

Appendix B: Igbo Disharmony Log..... 50

Appendix C: Tuvan VHC Screenshot..... 57

Appendix D: Igbo VHC Screenshot..... 58

## **Acknowledgements**

I must extend my deepest thanks to my advisor, Prof. K. David Harrison, without whom this thesis would never have been begun, much less finished. Prof. Vera Lee-Schoenfeld deserves credit for igniting my ongoing love for linguistics with her Intro class in freshman year, and she and the other wonderful professors at Swarthmore for encouraging and broadening my interest in the subject. I am also indebted to all of the linguists before me from whom I have learned and borrowed. My fellow Linguistics majors have been a source of much laughter and support, and I thank them for their comments and ideas. Lastly, I must thank my family for their unending love, support, and willingness to read my papers, no matter how esoteric the subject matter.

## **0. Abstract**

All languages have certain restrictions on what sounds are permitted, both in general and in combination with other sounds. A particular kind of constraint, vowel harmony, occurs in languages where the vowels within a word all belong to the same class ([±round], [±back], [±high], etc.), beyond the amount expected by statistical chance. Such languages are characterized by heavy affixation, wherein the vowels are underspecified and can change to suit the environment as needed. Like any pattern, however, there are exceptions. My goal in this thesis is to analyze two languages, Tuvan and Igbo, with respect to their systems of vowel harmony, and explore the ways that disharmony arises and is dealt with.

I start by outlining what vowel harmony is, what languages employ it, and what factors encourage and discourage its existence. In Section 2, I introduce Tuvan and provide a description of its phonology and morphology systems, as well as examples of disharmony, and Igbo receives the same treatment in Section 3. I compare and contrast the harmony levels of each language with Turkish, Japanese, and one another through statistical analysis of text corpora in Sections 4 through 7, and in Sections 8 through 10 I frame the results using constraints and Optimality Theory. I conclude with some musings on what the results could mean for our understanding of vowel harmony and the cognitive and learnability implications therein.

## **1. Introduction to Vowel Harmony**

Vowel harmony (henceforth VH) is a constraint by which “vowels agree with other vowels in terms of a particular phonological feature” (Harrison 2000: 111), such as backness, roundness, height, or Advanced Tongue Root (henceforth ATR). All vowels within a word or other morphological domain “look alike with respect to the active features” (Krämer 2003: 3). VH is typically present in languages with a rich morphological system, where affixes have varying vowels based on the surrounding sounds, and vowels within a root generally belong to the same harmony set (Casali 2008: 500).

There are two main ways that VH systems operate. The first, classified by Krämer as a morphologically-driven form of harmony (2003: 35), is known commonly

as root control harmony. In such a system, so-called "harmonizing affixes" (Casali 2008: 500) assimilate to roots, regardless of the vowel qualities in either morphological category (516). Clements & Sezer call this a "symmetrical" system (1982: 215), but in the interests of clarity, I will follow more recent literature and refer to it as the aforementioned root control harmony. Krämer presents a reputed case of affix-control harmony, in Fula (see also Paradis 1992 and Breedveld 1995), though previously Anderson (1980), McCarthy and Prince (1995) and Bakovic (2000) argued against the existence of such systems for various reasons. A resolution to the debate is outside the scope of this paper, and must be relegated to future studies.

The second type of VH is phonologically-driven (Krämer 2003: 35), and is generally called dominant-recessive harmony. In this type, which is rarely found outside of ATR harmony systems (Casali 2008: 514), one class of vowels (generally [+ATR]), if present in either root or affix, causes all of the other vowels within the word to assimilate to its value. Several Nilo-Saharan languages behave thusly, including Maasai, Bongo, and Lugbara (Linguist List). The only way [-ATR] vowels are pronounced is if none of the morphemes within the word contain a [+ATR] vowel underlyingly. In Clements' classification, this type of harmony is referred to as "asymmetrical," but again, in the interests of clarity, I will maintain the dominant (pun completely intended) nomenclature.

Directionality in VH is related to the categories above, and raises some interesting questions. In dominant-recessive systems, the issue is moot, since the

deciding factor is the presence of a dominant vowel, regardless of where it appears. However, most harmony systems belong to the root control subset, and thus any complete account must address the issue. Most VH processes spread the vowel values left-to-right (Finley 2008: 9, Clements & Sezer 1982: 219), though there are a non-negligible number of cases where right-to-left spreading occurs. While I will not be discussing directionality in great depth or utilizing constraints to describe the behavior in Tuvan and Igbo, I will briefly touch upon their respective directionality patterns (see Sections 2 and 3).

### **1.1 Languages that have VH**

VH is a relatively restricted phonological property, but it manifests in a number of diverse language families. Backness harmony, where vowels within a given harmonic domain are all either [+back] or all [-back], is present predominantly in the Eurasian Uralic and Altaic families (including the Altaic language Tuvan, which will serve as a case study for this paper) (Britannica). There are some African languages that exhibit a possible form of backness harmony, but it's questionable whether these cases are indeed motivated by backness harmony, or if they are examples of rounding harmony (Krämer 2003: 11). The latter, also called labial harmony, is characterized by the requirement that certain vowels agree in the feature [ $\pm$ round], and is present in many Altaic languages, including Tuvan. Rounding harmony is often restricted, and only applies when the affected vowel happens to “agree with respect to a second

feature like height or backness” (Krämer 2003: 7). The combination of two harmony systems, where vowels must agree in at least two feature domains as is the case in Tuvan, is called multiple feature harmony (Krämer 2003: 9). VH often applies only if the target and/or trigger possess certain feature qualities (15); in the case of Tuvan, rounding harmony only applies when the trigger vowel is [+round] and the affected vowel is [+high] (Kaun 1994, Harrison 2000).

Another common kind of VH pertains to the feature known as Advanced Tongue Root [ATR], where the relevant distinction is between tense [+ATR] and lax [-ATR] vowels. ATR harmony appears (almost?) exclusively in African languages, particularly in the Kwa, Cushitic, and Nilo-Saharan language families (Britannica). Igbo, a member of the first and the other language for case study in this thesis, exhibits a robust form of ATR harmony, explained further in Sections 3 and 6.

VH systems sometimes include vowels that do not participate in the harmony patterns. Such "neutral" vowels often arise from "gaps" in the vowel system; an example is taken from Finley (2008: 7):

If one [+HIGH] vowel does not have a [-HIGH] counterpart in the segmental inventory, that vowel cannot undergo [height] harmony.

A neutral vowel can be of two types, opaque and transparent. Vowels of the former class "start a new harmonic domain with their own feature specification" (Krämer 2003: 27), effectively blocking the spread of the harmonic feature value. On the other hand, transparent vowels allow the feature to spread through the vowel to the other side,



skipping the vowel itself but leaving the harmony pattern unchanged (Finley 2008: 7). Transparent vowels, as well, one would assume, as opaque vowels, can be always neutral, neutral in only certain environments, or unpredictably neutral (Kiparsky & Pajusalu 2003: 221). Neutral vowels of both types must "receive their feature values by independent specification" (Clements & Sezer 1982: 218). In both of my languages of focus, there are no neutral vowels (see Sections 2 and 3 for more detail), but some example languages that do contain such vowels are Finnish, Akan, and Mongolian (respective members of the Uralic, Kwa, and Altaic language families).

## **1.2 Reasons for VH**

It was previously thought that VH systems are most likely to arise in languages that have symmetrical vowel inventories (cf. Trubetzkoy 1969, cited in Harrison et al 2002: 3). Such inventories are examples of feature economy, which "favours maximising the number of phonemes that can be obtained by the free combination of a given set of features" (Clements 2003: 291). Tuvan has a near-perfectly symmetrical vowel system (see Section 2 for more), where the combination of [ $\pm$  high], [ $\pm$  back], and [ $\pm$  round], if used in every possible combination, gives  $2^3$ , or 8 possible vowels, every one of which is present in the Tuvan language. More recently, Harrison et al computed that harmony systems are actually likelier to arise in asymmetrical systems (2002: 5, 7), and I recommend the reader to their paper for a more thorough analysis.

Given a favorable vowel inventory, there are additional motivations for the

development of VH. One of the foremost theories is that it arises from co-articulation (Harrison et al 2002: 3), where the sounds of a word assimilate to one another. Having all the vowels in a word belong to the same class simplifies the cognitive and physical articulatory processes of the speaker and hearer. Once the first vowel is known, the possibilities for the remaining vowels within the word drops by at least half, something known as transitional probability. For the speaker, having the vowels occur in the same general area of the vowel space requires less effort, and is thus typologically preferable (Krämer 2003: 26). On the other side of the communication stream is the theory that a systematic harmony pattern can “enhance the probability that a given contrast or set of contrasts will be accurately perceived” by the hearer by “extending [the  $[\pm F]$ 's] duration” (Kaun 1995: 78). In a system with similar-sounding vowels, a harmony system can help differentiate between them (Kaun 1995: 78). In other systems, positional neutralization (Steriade 1993, cited in Kaun 1995) serves a similar purpose by restricting the distribution of perceptually subtle differences to places where they will be more easily identified. Similarly, Suomi argues that harmony serves to make weaker vowels more contextually predictable (1993, cited in Kaun 1995: 80). Another possible explanation for the existence of VH is its ability to indicate word boundaries, which, along with word stress, is helpful to the hearer in parsing morphological boundaries (Vroomen, Tuomainen, and de Gelder 1998).

All of these factors contribute to the development of a vowel harmonic system, but there are naturally factors that work in favor of the opposite, and diminish VH.

### 1.3 VH System Change

VH systems are not static. Old Turkic had a harmony level of near 100% (VHC, Swarthmore College), but its modern descendents range from disharmonic (Uzbek) to ~70% (modern Turkish) to 96% (Tuvan). One source of disharmony is loanwords. According to Harrison et al 2002, "the degree to which loanwords are mutated to be harmonic provides a possible diagnostic for the state of the harmony system" (5). Usually loanwords are introduced by a small number of bilingual speakers, and after being adopted by monolingual speakers are either modified to the harmony system or not (Harrison et al 2002: 8). Contrary to what one might think, the pronunciation of the loanword is not dependent on whether one speaks the source language or not, but rather on what the generally-accepted phonological form is (Poplack & Sankoff 1984: 105). One would assume from this, then, that the larger the proportion of bilingual speakers in the population, the more likely that the loanword would not be modified, though the question then arises whether the word in question is being used as a loanword or is actually a codeswitch. For more analysis of this, see Poplack & Sankoff 1984. For a more thorough exploration of the decay of VH, I recommend the reader to Harrison et al 2002, wherein the authors point to a number of internal and external factors that can bring about change in a harmony system, including the effects of vowel merger, wherein previously distinctive vowels in a language become phonemically identical.

Having given a rough theoretical background of VH and an examination of the contributing factors and some of the hurdles for the continuance of a VH system, I now turn to two specific case studies, Tuvan and Igbo.

## 2. Tuvan

Tuvan is a Turkic language spoken by approximately 250,000 people in Siberia, Mongolia, and China (Tuvan, Ethnologue). As previously mentioned, it has a symmetrical 8-vowel system, as shown below:

Table 2.1 Tuvan Vowel Distribution

i / y	ɯ / u
e / ø	ɑ / o

(Note: I have chosen to use IPA symbols rather than the letters traditionally used by Turkologists in order that the orthography within this paper remain consistent.)

Tuvan also has phonemic vowel length, so all of the vowels above surface as either short or long (Harrison 2000).

Tuvan is a heavily affixing language. Vowels within a word, including within affixes, must all share the same backness [ $\pm$ back] feature, i.e., must all be in the front class [i, y, e, ø] or in the back class [ɯ, u, a, o]. High vowels that follow a rounded vowel [y, ø, u, o] must also be [+round] (Harrison 2000: 111). Every vowel has a “harmonic counterpart” (Casali 2008: 500), a vowel differing in only one feature specification, as seen below:

Table 2.2 Vowel Feature Distribution

	i	y	e	ø	ɯ	u	ɑ	o
[high]	+	+	-	-	+	+	-	-
[back]	-	-	-	-	+	+	+	+
[round]	-	+	-	+	-	+	-	+

It is worth noting, however, that [ɑ] is lower than the other vowels in the [-high] class, which are described as being “slightly raised from a completely low position” (Krueger 95), and further back than the other [+back] vowels. In spite of this, however, [ɑ] is still the [+back, -round] member of the [-high] class, and fully partakes in harmony. This paper is concerned solely with backness harmony, and I refer the reader to Kaun 1995 for a detailed analysis of various types of rounding harmony within Turkic languages, and Harrison 2000 for Tuvan in particular.

Almost all suffixes in Tuvan (see section below for exceptions) are underlyingly unspecified; that is, the vowels in the affix carry limited feature values. An example suffix is the plural marker /-LAr/, which alternates as follows:

Table 2.3 Plural suffix alternation in Tuvan

#	Class	Noun stem	Genitive	Ungrammatical	English
1 a	[+back]	[orun]	[orun- <b>nar</b> ]	*[orun-ner]	‘Beds’
1 b		[maslo]	[maslo- <b>lar</b> ]	*[maslo-ler]	‘Oils’

1 c		[kuduu]	[kuduu- <u>lar</u> ]	*[kuduu-ler]	‘Ends’
1 d		[buga]	[buga- <u>lar</u> ]	*[buga-ler]	‘Oxen’
2 a	[-back]	[ygy]	[ygy- <u>ler</u> ]	*[ygy-lar]	‘Owls’
2 b		[børt]	[børt- <u>ter</u> ]	*[børt-tar]	‘Hats’
2 c		[jivi]	[jivi- <u>ler</u> ]	*[jivi-lar]	‘Firs’
2 d		[teve]	[teve- <u>ler</u> ]	*[teve-lar]	‘Camels’

(Vocabulary taken from Tuvan Talking Dictionary, Swarthmore College.)

From the data above, we can see that the suffix has two variants (not including consonant variation, which is explored thoroughly in Harrison 2000), one with [ɑ] and the other with [e]. The former attaches to stems with back vowels, and the latter to those with front vowels. The fact that the plural morpheme alternates is due to being *underspecified*, defined by Inkelas as a property of “a segment which surfaces with some phonological material M [that] is not specified for M in the input” (1994: 1). In this case, the plural morpheme is specified [-high] for height, but [Ø] for backness. This will become relevant in Section 9, for an Optimality Theoretic account of Tuvan. VH in Tuvan spreads from left to right; that is, the value for a feature is specified in the leftmost vowel of the root and spreads rightward, filling in the gaps in the underspecified vowel and providing the appropriate value for M.

## 2.1 Disharmony in Tuvan

While Tuvan does have remarkably robust systems of harmony, they are not without exceptions. Compound words can be back-disharmonic, often due to the

negative marker –čok [tʃok] (Harrison 2000: 114). Disharmony can also arise morphologically, through four non-harmonizing suffixes (allative, diminutive, durative, and sequential), and in some dialects of Tuvan, through the process of ablaut, which is used to communicate the intensive degree (115).

The other main source of disharmonic words are loanwords. Earlier-introduced words in Tuvan underwent vowel harmonization, so many of the Mongolian and older Russian words are harmonic. After bilingualism became more commonplace in the 1950s, the impetus to harmonize the words was significantly weakened. As a result, most recent English and Russian borrowings remain in their original disharmonic form (115-116). They are still, however, "subject to cluster simplification by means of consonant deletion or vowel epenthesis" (112), the latter of which is always high and potentially subject to rounding harmony.

Table 2.4 Exceptions in Tuvan

#	Source	Examples	English
1 a	Loan words	[maʃina]	‘machine’
1 b		[politika]	‘politics’
2 a	Non-alternating suffix	[teve-ma:]	‘camel-DIMINUTIVE’
2 b		[buga-ma:]	‘ox-DIMINUTIVE’

### 3. Igbo

Igbo, also spelled Ibo, is a Kwa language spoken by approximately 18 million people in Nigeria (Igbo, Ethnologue). As previously mentioned, it has an ATR VH

system, composed of two equal classes, tense ([+ATR]) and lax ([-ATR]). In the former class belong [i, e, u, o], and in the latter [ɪ, a, ʊ, ɔ]. Ihiunu and Kenstowicz use the feature [constr ph], for "constricted pharynx", as the distinguishing factor rather than [ATR] to better represent that "the [ɪ, a, ʊ, ɔ] set is marked in comparison to [i, u, e, o]" (1994: 1), but I have chosen to use the more common [ATR] in this analysis. Zsiga illustrates the vowel system by three feature classes, [±high], [±round], and [±ATR] (modified from 1997: 232):

**Table 3.1 Igbo Vowel Distribution**

	i	ɪ	e	a	u	ʊ	o	ɔ
[high]	+	+	-	-	+	+	-	-
[round]	-	-	-	-	+	+	+	+
[ATR]	+	-	+	-	+	-	+	-

In some western dialects of Igbo, the vowel system actually contains 9 vowels, with “an additional midfront vowel as the [-ATR] counterpart of [e], and in which the low vowel [is] transparent to harmony” (Welmers 1973, cited in Zsiga 1997: 232), but for the purposes of my investigation, the vastly more common 8-vowel system will be assumed. Ikekeonwu (1982) incorrectly considers [a] to be a [+ATR] vowel, but Zsiga explains that while [a] is both lower and further back than [e], it still operates as the [-ATR] counterpart to the latter (1997: 232), which is further backed by Casali (2008), who argues that [a] shows the same phonological markers (specifically voice quality) as the other [-ATR] vowels do (529). Thus, in Igbo as in Tuvan, every vowel has a harmonic counterpart, and there exist no neutral vowels.



VH can extend beyond the prosodic unit; one such example is the 3<sup>rd</sup> person pronoun (Carrell 1970: 5), which alternates between *o* [o] and *o̥* [ɔ̥] depending on the tenseness “quality of the vowel in the following verb stem” (2). This dependency on outside specification indicates that the pronoun, while written as a separate entity, is in fact at least an enclitic, and more likely an affix. VH in Igbo, unlike in Tuvan, is bidirectional; the vowel in the stem is specified for [ATR], and spreads rightward to suffixes and leftward to prefixes and pronouns.

### **3.1 Disharmony in Igbo**

The VH system in Igbo has very few "outright violations", both in "unmixed Igbo stretches [as well as] lone English-origin verbs" (Eke 1998: 192). According to Zsiga, "the only disharmonic morphemes are a few nouns beginning with [a]" (1997: 232), which she hypothesizes is leftover from the older 9-vowel system still present in the western Igbo dialects mentioned above. There are, however, several affixes that are written as though they do not alternate even though in spoken Igbo they usually do, and as is true in many vowel harmonic languages (Archangeli & Pulleyblank 1994: 3), compound words in Igbo, sometimes indicated by a hyphen, do not have to undergo harmony processes; I discuss my approach to these in Section 6.

## **4. Analysis**

Having given a basic outline of the phonology and VH systems of Tuvan and Igbo, I now turn to two forms of analysis: statistical and Optimality Theoretical.

## 4.1 Vowel Harmony Calculator

The Vowel Harmony Calculator (henceforth VHC) is a computer program developed by Prof. K. David Harrison (Swarthmore) and his students to analyze the level of vowel harmony within a given language, from a source corpus of data. It takes into account the vowel distribution of the source language and the average syllable count of polysyllabic words, from which it works out a Harmony Threshold, “what percentage of words would be harmonic by chance alone” (Harrison, O’Keefe, and Thomforde 2004). It then calculates the Harmony Index, which is the “percentage of harmonic words minus the harmony threshold” (VHC). A higher Harmony Index indicates a more harmonic language.

To use the VHC, a corpus of linguistic data must be collected and formatted (only ASC-II characters, one word per line, simple text) before being uploaded. The user can define whether there are any neutral vowels, if there are long vowels present (as is the case for both Tuvan and Igbo), and whether or not to simplify diphthongs (otherwise, the two vowels are counted as separate syllables). The relevant vowel classes are labeled and the vowels separated into their respective classes, and once all of the relevant data has been input, the VHC analyzes the data and gives back three log files of data, as well as a summary screen.

### Figure 4.1 VHC Screenshot

The summary screen gives the harmony threshold, harmony level, harmony

index, disharmony level, vowel class distribution, and average syllable number. The first log file is the Harmony Log, which contains essentially the same information as is present in the summary screen, with the addition of a breakdown of the number of syllables each word has. The second is the Frequency Log, which summarizes how frequently every vowel appears (in total, for only short vowels, only long vowels, and in what CV combinations). The final log file is the Disharmony Log, which is a list of all of the disharmonic words in the corpus. The summary screens for Tuvan and Igbo can be seen in Appendix C and D, respectively.

Larger corpora are preferable to smaller, as they more accurately represent the patterns and distribution of the language in question. Especially for phonological patterns like VH, it's hard to know whether results are statistically significant with a smaller corpus, as it's possible that the tokens just happened to exhibit signs of a pattern when there really isn't one, or conversely, that the tokens are skewed with exceptions to the rule, which conceals the presence of an actual pattern. For the same reason, it's good to have data from a variety of sources (books, Bible translations, newspapers, online articles, fables, etc.), in order that quirks in style or usage in one source form are diminished or revealed by the other sources.

## **5. Tuvan Data**

For the Tuvan corpus, I used data previously collected by Prof. Harrison and the University of Helsinki Language Corpus Project, and supplemented it with Tuvan

stories, songs, and articles I found online (for a complete list, see Data Sources). Tuvan is officially written in a modified Cyrillic script, but it has a couple of commonly-used English transliteration systems. Since the VHC only takes ASC-II characters, I replaced the special English characters with “standard” elements, as shown in the table below. The documents I found employed a couple of different systems, signified below by a slash between variants:

**Table 5.1 Tuvan Vowel Correspondences**

Tuvan	i	ü	e	ö	ï/y	u	o	a
IPA	i	y	e	ø	ɯ	u	o	ɑ
Corpus	i	U	e	O	I	u	o	a

## 5.2 Tuvan VHC results

Altogether, I ran 9,810 words, of which 7,763 were multisyllabic. Out of these 7,763, 7,390 were harmonic, giving a Harmony Level of ~95%. The Harmony Threshold, the percentage of words expected to be harmonic based on chance alone, was calculated to be ~37%. Therefore, the Harmony Index for the sample was ~58%. The distribution of vowels between the two classes was fairly skewed, with ~62% belonging to the back class and only ~38% to the front. This skewedness is taken into account by the VHC in calculating the Harmony Threshold and thus helps determine the Harmony Index. A summary of the output is shown below:

**Table 5.2 Tuvan VHC Results Summary**

Harmony Threshold: 37.51%  
 Harmony Level: 95.20%  
 Harmony Index: 57.68%

Front Vowels: 38.01%

Back Vowels: 61.99%

When we look at the disharmony log, consisting of 373 words, or 4.8% of the corpus, we can immediately see that a significant portion come from loanwords (as I don't speak Russian, I can only easily recognize the English-origin words, but I imagine that the number of disharmonic loanwords is actually quite a bit higher when Russian words are included). A table containing some of the disharmonic English-origin words (in transliteration) is below:

Table 5.3 Disharmonic Words in Tuvan

Tuvan	English		Tuvan	English
dialekt	dialect		teoriya	theory
taksi	taxi		kabinet	office
informatsiya	information		million	million
literatura	literature		partiya	party

See Appendix A for the complete disharmony log.

Using data previously obtained and analyzed by Prof. Harrison, I ran a corpus of ~10,000 words in Old Turkic to see how it compared to one of its descendents. The Harmony Level was a remarkable 96.19%, and the Harmony Index was 63.08%. The Disharmony Log, while obviously absent of any English-origin words, does nonetheless contain some recognizable Sanskrit terms, acquired via contact with Buddhists, such as *nirvan* “nirvana” and *bodisatv* “bodhisattva”.

I did the same for modern-day Turkish, and a corpus of ~44,000 words,

comprised of Bible translation and newspaper articles, contained only 61.79% harmonic words. This gives a Harmony Index of 32.94%, surprisingly low for the language often cited as being a prototypical example of VH. While some of the words in the disharmony log are recognizably loanwords (and again, there are doubtlessly many more than the ones I could recognize, given Turkey's exposure to Arabic), there are also many that possess Turkish-origin disharmonic roots. Examples of both are shown below:

Table 5.4 Disharmonic Words in Turkish

Borrowing?	Turkish	English
Y	kolesterol	cholesterol
Y	milyon	million
Y	akseptans	acceptance
Y	füzyon	fusion
N	fiyat	price
N	lisan	language
N	düşman	enemy
N	dünya	world

Clements and Sezer (1982) explain that Turkish roots, unlike those in Tuvan, often have two underlyingly-specified vowels which may not necessarily belong to the same class, which certainly contributes to disharmony.

So, while Tuvan is less harmonic than its progenitor, it is still significantly more harmonic than Turkish, which is itself much more harmonic than Japanese, which has an Harmony Index of 3.30% (Harrison, O'Keefe, & Thomforde 2004: Results). We can

already see from these four languages that the degree to which a language is VH is just that: a gradient. There are more-harmonic languages, and less-harmonic languages, and the question arises of where the cut-off should lie. Is a language with an Harmony Index of 10% harmonic? 15%? 20%? These are matters worth pursuing in further studies.

## 6. Igbo Data

I chose to investigate Igbo for a couple of reasons. Firstly, it exhibits a different kind of harmony from Tuvan, [ $\pm$ ATR] rather than [ $\pm$ back] and [ $\pm$ round], and I was interested to see how the harmony level compares. Secondly, it has a very precise orthography, which differentiates between all 8 vowels, necessary for building as accurate a corpus as possible (although this system is not always used, which I explain further below). Thirdly, though it is spoken by a significant population of Nigeria, there has been relatively little linguistic research done on it, and I hoped to contribute in some small way to the corpus of available information on the language.

To collect a corpus of Igbo data, I first turned to the internet. The official orthography of Igbo is in a modified Latin script, and differentiates between tense and lax vowels, with the latter being marked by a dot under the letter (with the exception of *a*, which is the lax counterpart to *e*). The text may optionally have tones marked as well, which, for the purposes of this paper, were disregarded. Not all written material I found included the dotted characters, which were necessary to the construction of an

accurate corpus, and thus any such material was excluded. As previously mentioned, Igbo has a fair number of compound words, which are sometimes disharmonic, but for the purposes of this investigation, I treated them as one word, as it seems likelier that they are viewed as such by the speakers.

The correspondences between the special characters used in Igbo, the IPA equivalencies, and the letters used in the corpus data are shown below:

Table 6.1 Igbo Vowel Correspondences

Igbo	i	ị	e	a	u	ụ	o	ọ
IPA	i	ɪ	e	a	u	ʊ	o	ɔ
Corpus	i	I	e	a	u	U	o	O

### 6.1 Igbo VHC Results

The corpus I collected consisted of 8,656 words, of which 5,252 were multisyllabic. Of the multisyllabic words, 4402 were harmonic, making the Harmony Level ~84%. As the Harmony Threshold was ~37%, the Harmony Index was ~46%. The vowel distribution was a little more even in Igbo than in Tuvan, with 44% belonging to the [+ATR] tense class, and 56% to the [-ATR] lax class. A summary of the results is shown below:

Table 6.2 Igbo VHC Results Summary

Harmony Threshold: 37.05%

Harmony Level: 83.82%

Harmony Index: 46.76%

Tense Vowels: 44.00%

Lax Vowels: 56.00%



A look at the disharmony log shows a fair number of English borrowings, including *makroni* “macaroni”, *amerika* “America”, and *kristian* “Christian”. The Bible excerpts contribute disproportionately to the disharmony, as *jizos* and *kraist* are both disharmonic and very frequent. The disharmonic Igbo-origin words were due either to a disharmonic root, or a harmonic root with a disharmonic affix. Examples of both are following (data taken from Williamson 1976):

Table 6.3 Disharmonic Words in Igbo

Disharmony source	Igbo	English
Root	kweṣi-ri	deserve-PAST
Root	banye-re	enter-PAST
Affix	bu-la	carry-PERF/AFFIRM
Affix	ke-ghị	tie-INDIC/NEG

I was curious as to the extent to which non-alternating affixes have an effect on the VH level. Using Ememanjo's Igbo grammar as my guide, I found two common suffixes that, while pronounced according to VH rules, are written as non-alternating. The first is *-ghị*, which is the inflectional indicative negative marker (Ememanjo 18), and the second is *-la*, which is used as the inflectional perfective affirmative marker, as well as the negative imperative suffix. I removed these suffixes from my data; in order to try to reduce the risk of removing segments that were actually part of roots, not suffixes, I only deleted those segments at the far-right edge of the word. This may have left some tokens with the suffixes still included, but I decided that was preferable to cutting out words that were, in fact, disharmonic.

The results are shown below:

Table 6.4 Igbo VHC Results Summary, Minus Disharmonic Suffixes

Harmony Threshold: 37.06%

Harmony Level: 84.25%

Harmony Index: 47.19%

Tense Vowels: 44.69%

Lax Vowels: 55.31%

The corpus was slightly smaller, reflecting the fact that removing the suffix from a disyllabic word made it monosyllabic and thus not a candidate for VH analysis. The vowel distribution also changed a little; the lax vowels ceded .69% to the tense vowels, which also makes sense, since the suffixes removed both contained lax vowels.

I was also interested in seeing as to whether the form of the source of the text would change the harmony value, so I collected a few printed books in Igbo and scanned and uploaded them in the hope of running them through an Optical Character Recognition program so that I could turn them into plain text and run them through the VHC. Unfortunately, the diacritics used in Igbo are not supported or recognized by the program available to me (ABBY), and since they are contrastive in the language, any output data that lacked them was not ideal. Instead, I typed up selections from a couple of the texts to build a corpus, albeit one that is much smaller (only 2,719 multisyllabic tokens) than I had initially planned. The results can be seen below:

Table 6.5 Igbo Print Sources VHC Results Summary

Harmony Threshold: 39.79%

Harmony Level: 87.13%

Harmony Index: 47.34%

Tense Vowels: 40.75%

Lax Vowels: 59.25%

The values for the various harmony measurements are quite similar to those for the internet sources; interestingly, the vowel class is more skewed, with a larger percentage belonging to the lax class. This increases the probability of harmony occurring, reflected in the slightly higher Harmony Threshold, and the Harmony Level is correspondingly higher, but the difference between the two remains nearly identical as that in the previous corpus. This perfectly illustrates the value of the Harmony Index—it effectively negates the effects that vowel class skewing may present and gives a less-biased account of the harmony level in a language. I cannot explain why the vowel classes in the two samples aren't more similar, though I imagine the smallish corpus sizes do contribute somewhat.

In the interests of thoroughness, I removed the same two suffixes as before and received back the following:

Table 6.6 Igbo Print Sources VHC Results Summary, Minus Disharmonic Suffixes

Harmony Threshold: 39.83%

Harmony Level: 87.94%

Harmony Index: 48.11%

Tense Vowels: 41.22%

Lax Vowels: 58.78%

Comparably to the previous results, the Harmony Index increases a little, but only by 0.77%. The number of tokens decreased as well, dropping to 2,679, and the vowel classes became a little more equal, with the removal of the lax suffixes. The difference between the Harmony Indexes of the print and web sources with the disharmonic affixes is proportional to the difference between them without, suggesting that at least those two disharmonic affixes appear equally as frequently in both mediums.

It is worth noting that the actual harmony level in Igbo is higher than can be accounted for by the VHC. The pronouns, as mentioned in Section 3, do behave harmonically, but as they are written separately from the verb with which they harmonize, the VHC has no way of recognizing this fact. Perhaps some future researcher, more gifted in computer science than myself, can find a way to analyze the separate pronouns with the associated verb.

There has not yet, to my knowledge, been a similar analysis of any other ATR-harmony languages, so I can't say how the Igbo results compare to others of its ilk. I hope that future researchers will be able to fill in the gaps in current documentation and create a more fully representational account of VH.

## **7. Statistical Summary**

Both Tuvan and Igbo show high Harmony Indexes (> 45%), though the forms

that their VH takes differ. As previously mentioned, it is an ongoing question of where the boundary between “harmonic” and “not harmonic” languages lies, and I do not currently feel equipped to try to provide an answer for it. We can say that both Igbo and Tuvan are strongly harmonic, as, to a lesser extent, is Turkish, while Spanish is clearly not harmonic. The degree to which the Harmony Index is explanatory versus descriptive is also something to consider, but regardless, it is a productive and informative tool for expanding our understanding of VH.

Now that I've attempted to elucidate the degree and type of VH in Tuvan and Igbo, I turn to a theoretical framework through which we can view VH patterns.

## **8. Optimality Theory and Harmony**

Optimality Theory (henceforth OT), first presented by Prince & Smolensky in 1993, is a paradigm under which language patterns can be explained by the interaction of constraints. Kirchner summarizes it rather more succinctly than I could, saying:

"The phonological component of the grammar consists solely of a set of universal, violable constraints, ranked on a (partially) language-specific basis" (Kirchner 1993: 1).

OT has three main components. The first, GEN, short for "generator", "takes an underlying representation ('input') and returns a (possibly infinite) set of possible surface forms ('output candidates')" (Zuraw 2000: 15). Richness of the base is the nomenclature used to refer to the assumption that "there are no language-specific restrictions on inputs" (McCarthy 2005: 8). EVAL, short for "evaluator", takes the

output candidates and "chooses the candidate that best satisfies a set of ranked constraints; this optimal candidate becomes the surface representation" (Zuraw 2000: 15-16). CON is the set of universal constraints, which all languages share but rank differently (McCarthy 2005: 8); the order the constraints are ranked determines the "optimal" output candidate. The ranking order determines a harmonic ordering, called H-Eval (Tesar & Smolensky 2000: 24), of the outputs; when constraints conflict, the highest ranked one strictly dominates those lower, shown as  $A \gg B$ , where the higher ranked constraints are to the left, and "even perfect performance on B can't overcome inferior performance on A" (McCarthy 2005: 8), similar to how *az* will always precede *ba* in an alphabetic ordering.

OT constraints are of two types, *faithfulness* and *markedness*. The former requires that the output or observed form adhere as precisely as possible to the input or underlying form, and the latter demand that the output form be as phonologically unmarked as possible (Finley 3). The degree to which a segment is marked is determined by how perceptually salient, easy to pronounce, and simple to learn it is (Jakobson 1968, cited in Finley 2008). The conflict between these two types of constraints, as well as the specific constraints within each category, characterizes OT analysis.

### **8.1 Faithfulness Constraints**

Faithfulness constraints comprise two main families. The first, named PARSE,

requires that the material present in the underlying input form manifest in the output form and works to prevent the failure of “underlying material to be structurally analyzed” (Prince & Smolensky 1993: 25). A primary constraint in the PARSE family is MAX, which mandates against deletion (Finley 2008: 160).

1) MAX: No deletion.

The second family, called FILL, mandates that every node in the output must be properly filled, i.e., that information in the output be “strictly based on underlying material” (Prince & Smolensky 1993: 25); the corresponding constraint here is DEP, which is violated by epenthesis/insertion (Finley 2008: 161).

2) DEP: No insertion.

PARSE and FILL are combined in Finley's RECIPROCITY, which “is violated whenever the feature values at the projection and the pronunciation level are not the same” (158).

Another class of constraints is IDENT-IO, also called IDENT[F], which requires a segment to carry the same feature identities in the output as was in the input.

3) IDENT[F]: A [ $\alpha$ F] segment in the input must not have a [ $-\alpha$ F] correspondent in the output. (Kiparsky & Pajusalu 2003: 223).

This will prove useful for describing VH, especially for when it dominates markedness constraints, explained further below.

## 8.2 Markedness Constraints

As previously mentioned, markedness constraints work to make the output as

perceptually and articulatorily distinct and easy as possible. For VH, the relevant concern is featural agreement between segments within a word. Kaun explains that "from the perceptual standpoint, it is advantageous to extend the duration of *all* phonological features," which leads to the constraint EXTEND $\alpha$ :

4) EXTEND $\alpha$ : all features want to spread. (Kaun 100)

Kiparsky & Pajusalu use a similar constraint, which they call SPREAD following Padgett, which is

5) SPREAD[F]: If any segment is associated with F, then every segment is associated with F. (2003: 223)

Another approach used to describe VH is the AGREE constraint, which "requires adjacent vowels to share the same phonological feature value" (Finley 2008: 25):

6) AGREE[F]: Adjacent segments must have the same value of the feature [F].

(Bakovic, 2000; Lombardi, 1999; Pulleyblank, 2002, cited in Finley 2008: 25)

Others have used the ALIGN constraint to explain VH, but I feel that the EXTEND constraint is easier to use for my data and provides an acceptable analysis. This is especially true for the Igbo data, in which, as mentioned above, the pronouns receive their featural identity from outside the prosodic word.

In a different vein, Andrew Nevins has argued for a target-driven approach to harmony, rather than a donor-driven one. In essence, "harmony is a search initiated by a 'needy' vowel for the features that it requires" (Nevins 2010: 20). This contrasts to the previously-discussed methods, which involve an underlyingly-defined vowel "spreading" to all undefined vowels within its purview. Following his approach, the



relevant constraints would be along the lines of SEARCH and COPY (Nevins 2010: 20) rather than IDENT and EXTEND. Given the novelty of this approach, I will stick with the more traditional view for the purposes of this paper, but I do find that Nevins makes some valid points and may serve as a catalyst for further (re)analysis of harmony.

## 9. Tuvan OT Analysis

The goal here is to offer a possible ranking of constraints for backness harmony in Tuvan. Harrison 2000 offers a very complete analysis using ALIGN constraints, which can account for epenthesis vowels, but I am going to rely on AGREE constraints here instead for the sake of simplicity. I will also assume the existence of underspecified archiphonemes in my following analyses, as mentioned in Section 2, and refer the reader to Inkelas 1994 for an explanation of the feasibility of archiphonemes in OT.

So, the constraints I will call upon to explain Tuvan VH are as follows:

AGREE[back]: A markedness constraint that requires that all vowels within a word have the same backness value.

IDENT[back]: A faithfulness constraint that requires that all outputs have the same backness feature value as the corresponding input; for vowels that are underspecified, this constraint does not apply.

Backness VH in Tuvan can be roughly explained with these two constraints. The ranking order of the constraints can be either

a) IDENT[back] >> AGREE[back]

b) AGREE[back] >> IDENT[back]

Using the plural suffix as our first example, and attaching it to the front-voweled [teve] *teve* “camel”, we can see that either order gets us the correct, attested output:

Tableau 9.1

	/teve-LAr/	AGREE[back]	IDENT[back]
☞	[teveler]		
	[tevelar]	*!	

Tableau 9.2

	/teve-LAr/	IDENT[back]	AGREE[back]
☞	[teveler]		
	[tevelar]		*!

However, we can turn to a non-alternating morpheme, namely the [+back] negative suffix *-čok* [tʃok], in order to see what happens when the affix contains an underlyingly represented vowel.

If we use a stem that has back vowels, in this case [tuvar] *tyvar* “find”, we run into the same problem as with the plural suffix, in that either constraint ranking hierarchy gives the correct output, seen below:

Tableau 9.3

	/tuvar-tʃok/	AGREE[back]	IDENT[back]
☞	[tuvartʃok]		

Tableau 9.4

	/tuvar-tʃok/	IDENT[back]	AGREE[back]
☞	[tuvartʃok]		

If we choose a stem that contains front vowels, here *čür* [tʃi:r] “eat”, we can finally come to some conclusions:

Tableau 9.5

	/tʃi:r-tʃok/	AGREE[back]	IDENT[back]
	[tʃi:rtʃok]	*!	
☞☹	*[tʃi:rtʃek]		*

In the table above, ☹ indicates that the winning output is not the attested form, meaning that the constraint hierarchy does not accurately describe the data. If we switch the order, though, we see:

Tableau 9.6

	/tʃi:r-tʃok/	IDENT[back]	AGREE[back]
☞	[tʃi:rtʃok]		*
	*[tʃi:rtʃek]	*!	

This time the winning output is the correct option. We now know that the constraint ranking to derive Tuvan VH is

$$\text{PARSE[back]} \gg \text{AGREE[back]},$$

where the former constraint is active when the vowel is underlyingly specified for backness.

Disharmonic loan words are allowed to surface through the same ranking order, though as Harrison points out, this ranking order is relatively recent, as older loan words did undergo backness harmonization (2000: 117).

## 10. Igbo OT Analysis

AGREE[ATR] and IDENT[ATR] are the relevant constraints for Igbo. Modeling off of the conclusions from above, we can see if the same ranking order will give out the correct outputs. We start with an alternating prefix, namely /I-/, the infinitive marker.

I browsed Emamanjo's grammar and chose *si* [si] “cook” (1978: 18) for my tense vowel stem, whose tableau is shown below:

Tableau 10.1

	/I-si/	IDENT[ATR]	AGREE[ATR]
☞	[isi]		
	[i̇si]		*!

Fortunately, this ranking gives the correct output. Using a stem with a lax vowel, *sị* “say”, we see:

Tableau 10.2

	/I-sị/	IDENT[ATR]	AGREE[ATR]
	[isị]		*!
☞	[i̇sị]		

Again, we receive the attested form.

We now turn to the non-alternating indicative negative suffix /-ghị/ to confirm that the same ranking order gives an accurate result. Using the tense root *si*, we obtain the following tableau:

Tableau 10.3

	/si-ghi/	IDENT[ATR]	AGREE[ATR]
☞	[sighi]		*
	*[sighi]	*!	

And, praise the Flying Spaghetti Monster, the correct output is selected by the ranking order.

## 11. OT Summary

The constraints IDENT[F] and AGREE[F], where [F] is [round] in Tuvan and [ATR] in Igbo, suffice to describe the tokens, both harmonic and disharmonic, in the two languages when ranked as

$$\text{IDENT}[F] \gg \text{AGREE}[F].$$

## 12. Conclusion

Using the language of OT, we can say that speakers must learn which roots and affixes are exceptionally marked for the feature in question so as to know when to apply IDENT. This would seem to diminish the efficacy of a VH system, since part of the appeal, as discussed in Section 1.2, is the cognitive ease such a system brings. One must ask at what point the exceptions and disharmonic elements make the system no longer worthwhile. When do the words or morphemes marked for IDENT become so significant that the rules governing harmony patterns are no longer effective enough to warrant remembering?

The Harmony Index can provide a unique view of the degree to which a language employs IDENT over AGREE, and from that, the degree of harmony it exhibits. Both Tuvan and Igbo implement IDENT rarely, reflected in their high Harmony Indices, whereas Japanese always employs IDENT, which may fulfill AGREE incidentally, but the latter is likely so far down the list of mental constraints as to be completely irrelevant.

My analyses of Tuvan and Igbo are but a tiny piece of the VH puzzle. They are both obviously very harmonic, and can serve as examples of the upper bounds of the potency of VH systems. Turkish lies somewhat lower on the spectrum, but there is a big gap between it and Japanese in terms of degree of harmony. It is my hope that future studies will better fill in the picture of where harmonic-with-exceptions becomes non-harmonic, and what the divide means for learnability and cognitive processes.

## Bibliography

- Albright, Adam. "Lexical and morphological conditioning of paradigm gaps." MIT: 2006. Web.
- Archangeli, Diana and Douglas Pulleyblank. *Grounded Phonology*. Cambridge, MA: MIT Press, 1994.
- Bakovic, E. "Harmony, dominance and control." Diss. Rutgers University, 2000.
- Breedveld, J.O. *Form and Meaning in Fulfulde. A Morphological Study of Maasinankoore*. Leiden: CNWS Publications, 1995.
- Carrell, Patricia L. *A Transformational Grammar of Igbo*. Cambridge: University Press, 1970.
- Casali, Roderic F. "ATR Harmony in African Languages." *Language and Linguistics Compass* 2.3 (2008): 496-549. *JSTOR*. Web.
- Clements, G. N. "Feature Economy in Sound Systems." *Phonology* 20.3 (2003): 287-333.
- Clements, G. and E. Sezer. "Vowel and Consonant Disharmony in Turkish." *The Structure of Phonological Representations* (Part II). Ed. H. van der Hulst and N. Smith. Foris: Dordrecht, 1982. 213-255. Print.
- Eke, Ejike. "Lending Credence to a Borrowing Analysis: Lone English-Origin Incorporations in Igbo Discourse." *International Journal of Bilingualism* 2 (1998): 183-201. Web.
- Ememanjo, E. Nolue. *Elements of Modern Igbo Grammar: A Descriptive Approach*. Ibadan: Oxford University Press, 1978.
- Finley, Sara. "Formal and Cognitive Restrictions on Vowel Harmony." Diss. Johns Hopkins University, 2008. Web.
- Harrison, K. David. "Topics in the Phonology and Morphology of Tuvan." Diss. Yale University, 2000. Web.
- Harrison, K. David, Mark Dras, and Berk Kapicioglu. "Agent-Based Modeling of the Evolution of Vowel Harmony." *32nd NELS Conference*. New York: 2002. Web.
- Harrison, K. David, Michael O'Keefe, and Emily Thomforde. Vowel Harmony Calculator, computer program. Swarthmore College, 2004.
- Harrison, K. David, Michael O'Keefe, and Emily Thomforde. "Vowel Harmony: Results." [http://www.swarthmore.edu/SocSci/harmony/public\\_html/results\\_new.html](http://www.swarthmore.edu/SocSci/harmony/public_html/results_new.html). 2004. Accessed Oct. 10, 2010.

- Harrison, K. David and Eric Raimy. "Language as an Emergent System." *Soundings: An Interdisciplinary Journal* XC.1-2 (2007): 77-90. Web.
- Hayes, Bruce P. "Four Rules of Inference for Ranking Argumentation." UCLA: 2003. Web.
- Ihiunu, Peter, and Michael Kenstowicz. "Two Notes on Igbo Vowels." MIT: 1994. Web.
- Ikekeonwu, Clara I. "Borrowings and Neologisms in Igbo." *Anthropological Linguistics* 24.4 (1982): 480-486. *JSTOR*. Web.
- Inkelas, Sharon. "The consequences of Optimization for Underspecification." UC Berkeley: 1994. Web.
- Jakobson, R. *Child language, aphasia and phonological universals*. The Hague: Mouton, 1968.
- Kaun, Abigail Rhoades. "The Typology of Rounding Harmony: An Optimality Theoretic Approach." Diss. UCLA, 1995. Web.
- Kiparsky, Paul, and Karl Pajusalu. "Towards a typology of disharmony." *The Linguistic Review* 20 (2003): 217-241. Web.
- Kirchner, Robert. "Turkish Vowel Harmony and Disharmony: An Optimality Theoretic Account." *Rutgers Optimality Workshop I*: October 22, 1993. Web.
- Krämer, Martin. *Vowel Harmony and Correspondence Theory*. Berlin: Mouton de Gruyter, 2003. E-book.
- Lewis, M. Paul (ed.). *Ethnologue: Languages of the World, Sixteenth edition*. Dallas, TX: SIL International, 2009.
- Lombardi, L. "Positional faithfulness and voicing assimilation in optimality theory." *NLTT* 17 (1999): 267-302.
- Máire Ní Chiosáin with J. Padgett (2001) "Markedness, Segment Realization and Locality' in Spreading" In: Linda Lombardi, ed (eds). *Constraints and representations: segmental phonology in Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, John J. "Handout #1: Background, Origins, and Overview of OT." University of Massachusetts, Amherst: 2005. Web.
- Nevins, Andrew. *Locality in Vowel Harmony*. Cambridge, MA: MIT University Press, 2010.
- Paradis, Carole. *Lexical Phonology and Morphology: The Nominal Classes in Fula*. New



- York, London: Garland, 1992.
- Pater, Joe. "Learning a Stratified Grammar." *Proceedings of the 29th Boston University Conference on Language Development*, edited by Alejna Brugos, Manuella R. Clark-Cotton, and Seungwan Ha, Cascadilla Press, Somerville, MA. 2005. 482-492.
- Pater, Joe. "Exceptions in Optimality Theory: Typology and Learnability." Umass Amherst Master's Abstract, 2005.
- Poplack, Shana and David Sankoff. "Borrowing: the synchrony of integration." *Linguistics* 22 (1984): 99-135. Web.
- Prince, Alan, and Paul Smolensky. *Optimality Theory: Constraint Interaction in Generative Grammar*. ROA: 1993, 2002.
- Pulleyblank, D. "Harmony drivers: No disagreement allowed." *BLS* 28 (2002): 249-267.
- Steriade, Donca. "Positional neutralization." *NELS* 24. New York: 1993.
- Suomi, Kari. "Palatal harmony: A perceptually motivated phenomenon?" *Nordic Journal of Linguistics* 6 (1983): 1-35.
- Tesar, Bruce and Paul Smolensky. *Learnability in Optimality Theory*. Boston: MIT Press, 2000. E-book.
- Tuvan Talking Dictionary, Swarthmore College, 2010.  
<http://www.swarthmore.edu/SocSci/tuvan/dict/>.
- "vowel harmony." *Encyclopædia Britannica*. 2010. Encyclopædia Britannica Online.  
[http://www.swarthmore.edu/SocSci/harmony/public\\_html/vhc.html](http://www.swarthmore.edu/SocSci/harmony/public_html/vhc.html).
- Vroomen, Jean, Jyrki Tuomainen, and Beatrice de Gelder. "The Roles of Word Stress and Vowel Harmony in Speech Segmentation." *Journal of Memory and Language* 38 (1998): 133-149.
- Welmers, William E. *African Language Structures*. Los Angeles: University of California Press, 1973.
- Williamson, Kay. *Dictionary of Onicha Igbo*. Cambridge: Ethiope Press, 1972, 2006. Web.
- Zsiga, Elizabeth C. "Features, Gestures, and Igbo Vowels: An Approach to the Phonology-Phonetics Interface." *Language* 73.2 (June 1997): 227-274. JSTOR. Web.
- Zsiga, Elizabeth C. "A mismatch between morphological and prosodic domains:

evidence from two Igbo rules." *Phonology* 9 (1992): 101-135. Web.  
Zuraw, Kie Ross. "Patterned Exceptions in Phonology." Diss. UCLA, 2000.

### Data Sources

#### Tuvan Source List

Aa-Shuu-Dekei-Oo, <http://www.tyvawiki.org/wiki/Aa-Shuu-Dekei-Oo>.  
Amyd baryp, [http://www.tyvawiki.org/wiki/Amyd\\_baryp](http://www.tyvawiki.org/wiki/Amyd_baryp).  
Ancestors, <http://www.tyvawiki.org/wiki/Ancestors>.  
Anton Pavlovich Chekhov, [http://www.tyvawiki.org/wiki/Anton\\_Pavlovich\\_Chekhov](http://www.tyvawiki.org/wiki/Anton_Pavlovich_Chekhov)  
Bay-Tayga, <http://www.tyvawiki.org/wiki/Bay-Tayga>.  
Black Mountain, [http://www.tyvawiki.org/wiki/Black Mountain](http://www.tyvawiki.org/wiki/Black_Mountain).  
Buura, <http://www.tyvawiki.org/wiki/Buura>.  
Chashpy-Khem, <http://www.tyvawiki.org/wiki/Chashpy-Khem>.  
Chok-la kizhi yry, [http://www.tyvawiki.org/wiki/Chok-la kizhi yry](http://www.tyvawiki.org/wiki/Chok-la_kizhi_yry).  
Dadyr-dadyr, <http://www.tyvawiki.org/wiki/Dadyr-dadyr>.  
DaNgyna, <http://www.tyvawiki.org/wiki/Da%C5%8Bgyna>  
Dialects and the Literary Language,  
[http://www.tyvawiki.org/wiki/Dialects and the Literary Language](http://www.tyvawiki.org/wiki/Dialects_and_the_Literary_Language).  
Khajyrakan, <http://www.tyvawiki.org/wiki/Khajyrakan>.  
Milk Tea, [http://www.tyvawiki.org/wiki/Milk Tea](http://www.tyvawiki.org/wiki/Milk_Tea).  
Our Native Land, [http://www.tyvawiki.org/wiki/Our Native Land](http://www.tyvawiki.org/wiki/Our_Native_Land).  
Shaman Wolf and Scholar Wolf,  
[http://www.tyvawiki.org/wiki/Shaman Wolf and Scholar Wolf](http://www.tyvawiki.org/wiki/Shaman_Wolf_and_Scholar_Wolf).  
The Marmot and the Ram, [http://www.tyvawiki.org/wiki/The Marmot and the Ram](http://www.tyvawiki.org/wiki/The_Marmot_and_the_Ram).  
The Orphan Boy, [http://www.tyvawiki.org/wiki/The Orphan Boy](http://www.tyvawiki.org/wiki/The_Orphan_Boy).  
The Raven and the Owl, [http://www.tyvawiki.org/wiki/The Raven and the Owl](http://www.tyvawiki.org/wiki/The_Raven_and_the_Owl).  
The Smart Boy, [http://www.tyvawiki.org/wiki/The Smart Boy](http://www.tyvawiki.org/wiki/The_Smart_Boy).  
Throatsinging – Treasure of the Tuvan People,  
<http://www.tyvawiki.org/wiki/Throatsinging>  
[http://www.tyvawiki.org/wiki/Throatsinging  
%E2%80%93 Treasure of the Tuvan People](http://www.tyvawiki.org/wiki/Throatsinging_%E2%80%93_Treasure_of_the_Tuvan_People).  
Tuvan corpus, Swarthmore College, 2004.  
[http://www.swarthmore.edu/SocSci/harmony/public html/corpora/Tuvan.txt](http://www.swarthmore.edu/SocSci/harmony/public_html/corpora/Tuvan.txt).

### Igbo Source List: Web

Biafra, <http://ig.wikipedia.org/wiki/Biafra>

Egbuighi-egbu , [http://ig.wikipedia.org/wiki/Egbuighi\\_igbu](http://ig.wikipedia.org/wiki/Egbuighi_igbu)

Ì Choro İma nke Bù Eziokwu? [http://www.watchtower.org/ib/kt/article\\_01.htm](http://www.watchtower.org/ib/kt/article_01.htm)

Tower of Babel in Igbo, <http://www.omniglot.com/babel/igbo.htm>

<http://www.gotquestions.org/Igbo/Ndu-Ebighi-Ebi.html>

<http://www.gotquestions.org/Igbo/Natawo-Mgbaghara-mmehie.html>

<http://www.gotquestions.org/Igbo/Onyenzoputa-nke-onwe.html>

<http://www.gotquestions.org/Igbo/Atumatu-Nke-Nzoputa.html>

<http://www.ohchr.org/EN/UDHR/Pages/Language.aspx?LangID=igr>

### Igbo Source List: Print

Agwuna, Stella Ogechukwu. *Iko Onye Ji Ama...Akwukwo Ejije Igbo*. Onitsha: Lincel Publishers, 2006.

*Bible Nso*. Lagos: The Bible Society of Nigeria.

Osuagwu, B.I.N and E.C. Nwana. *Omeokachie Omenuko*. Umuahia, Abia State: Ark Publishers, 1999.

## Appendix A: Tuvan Disharmony Log

(Dial	buterbrod	kambe:t	muzey
(domak/sOs)	daStIkIZe	kanikul	myaCik
-tinnIn	da:rakCIzi	karZi	naklonenie
Ca:vey	dekabr	kartina	nakloneniezi
Cemodan	dialekt	kinZal	nareCie
Cesnok	duru`\e1a	kino	nasekom
Cudeksineri	e`\cadi	klimat	nedelya
Cudey	ekran	konfeta	ofis
InCeS	ekzamen	korzina	ofitser
IrakCe	epos	kreslo	ogurec
Respublika)	ergeZok	krolik	okean
Sa: = bile	fabrika	kude:	oliva
Sartila:	fakul'\edtet	kvartira	omonim
Sina	familiya	kvitantsiya	padeZ
Sveytsar	ferma	legenda	padeZi
Zaket	festival'	lektsiya	panedel'nik
a:lInCe	fevral'	lenta	papiros
ada:nCe	fizika	liCinka	parazit
adres	fotografiya	limon	partiya
aeroport	garCitsa	literatura	passaZir
akki	golovastik	literaturanIN	pi:va
alfavit	gorizont	literaturlug	piva
ami	gostinitsa	maSina	piyavka
anglile:r	gruzovik	magazin	planeta
apteka	ideal	magnito(f)on	plotina
arenda	ideya	matematika	poema
argi:r	idioma	material	poezd
armchair	informatsiya	menyu	poeziya
armiya	instrument	metall	politika
asbest	iyo:	metro	pomidor
avariya	iyul'	mifologiya	pressa
baZINInCe	iyun'	militsiya	problema
badIlaZi:Skin	kIrInCe	milyon	professiya
badIlaZir	kO:ru	minut(a)	provintsiya
basseyn	kaCeli	moroZenoe	pulemyot
biblioteka	kabinet	mototsikl	pyatnitsa

rabbit	turizm	o:ldarZe	teoriyalIg
radio	udarenie	uZuraskIZe	materialdarga
rakovina	udarenye	SayZe	literaturanIN
redaktorla:r	uley	kIrInCe	Cexov
reklama	univermag	sovet	Cexov
rekomentatsiya	universitet	million	gimnaziyanI
respublika	veSCestvo	million	universitediniN
restoran	viSnya	sovet	medicina
revolyutsiya	video	armiyaniN	fakultedinge
rifma	vinograd	revolyutsiyanIN	Cexov
rolik	vintovka	kommunistig	universiteti
samalyet	voskresenye	partiya	CexovtuN
sapo(o)Znik	vtornik	emgeZok	Cexov
sekretar'	xONnu	kommunistig	Cexov
sekta	xOIcok	partiyaniN	literaturlug
sektsiya	xUlumzUrU:SkUn	SartaygIZe	povest
sekunda	xaladi:lnik	argazInCe	garSin
sentyabr'	xure:	xunayeviCiniN	rossiyada
sezon	oglundCe	traktorist	Cexov
simvol	uruglarIm!be:r!be:r!	sanCleviC	saxalin
sizo	nazIla:Ze	literatura	ortulu:nCe
sliva	ok!kUske!kUske!	institudu	Cexov
sreda	xUUmzUreviSa:n	arxivinden	saxalin
stansiya	tir!baStaktanIp	xava:ngCe	saxalin
stsena	CeCekma:	xava:nCe	CexovtuN
student	boris	kIrIglZe	Cexov
tabletka	arInnarInCe	aNmeN	versta
taksi	boris	xovuzunCe	Cexov
telefon	aspa:nCe	da:nCe	Cexov
telegramma	er!salIp	dayanip	xanIzInCe
televizor	aparganinIN	o:loviCiniN	Cexovka
telo	arnInCe	o:loviC	CexovtuN
temperatura	nazIlagIZe	aziya	emnetpiSa:n
teoriya	men!kIrIp	geografitIg	Cexov
tetrad'	epCoksunup	aziyaga	masterstvonu
tomuya:ri	SimCe:rSInIm	evropaga	Cexov
turguZe	o:ldarImiStinde	amerikaga	teatrI
turist	katkiZe	natsiya	CexovtuN

teatrni	politiktig	literaturlug	leksika
teatrniN	social	dialektige	dialektide
scenazInga	ekonomiktig	dialektilerniN	dialektide
CexovtuN	nacional	literaturlug	grammatiktig
viSnyovIy	literaturlug	dialektiler	dialektide
Cexov	dialektilerge	dialektileriniN	dialektide
gorkiy	literaturlug	sistemazI	dialektide
Cexov	dialektilerniN	dialektiler	dialektilerde
gorkiyiniN	literaturlug	dialektileriniN	barzImze
Cexov	literaturlug	sistemazI	dialektide
akademiyazI	dialektiniN	dialektilerden	dialektide
maksim	dialektizinge	dialekt	dialektide
gorkiyini	literaturlug	dialekt	dialektide
policiyaniN	dialektilerniN	dialekt	dialektilerde
proletariy	dialektiziniN	dialekt	dialektide
revolyusCu	grammatiktig	dialekt	dialektileriN
gorkiyini	literaturlug	dialekt	sistemazInIN
nikolayga	grammatiktig	dialekt	sxemazI
rezolyuciya	dialektilerniN	dialekt	dialekt
akademikterniN	dialektizmner	dialekt	fonetikazI
vladimir	dialektiniN	dialekt	leksikazI
korolenko	institutka	fonetikazI	grammatikazI
Cexov	dialektiziniN	leksikazI	literaturlug
akademik	literaturlug	grammatikazI	dialektige
CexovtuN	dIlCe	literaturlug	literaturlug
iyul	dialektilerniN	dialekt	dialektiden
Cexov	dekabr	dialekt	literaturlug
Cexov	dialekt	dialekt	nacional
CexovtuN	dialektilerni	basseyneri	literaturlug
literaturlug	nakloneniN	dialektilerniN	dialektiler
dialektilerniN	CiZe:le:rgebarzImize	oCerkke	literaturlug
dialektilerniN	barzImze	dialektilerniN	dIlCe
samodiy	barzINize	dialektide	dialektilerden
feodaldIg	barzINze	dialektide	dialekt
administratsiya	literaturlug	dialektide	dialektilerden
dialektilerniN	leksikazI	dialektide	enesajIm
nacional	dialekt	dialektide	daSce
revolyutsiya	dialektiniN	dialektilerde	daSce

CaZINajenIp	gorizont	nareCie	sekunda
(Dial	gostinitsa	nasekom	sentyabr'
-tinnIn	informatsiya	ofis	sezon
Ca:vey	instrument	ofitser	simvol
Cemodan	iyo:	okean	sizo
Cesnok	iyul'	oliva	sliva
Cudeksineri	iyun'	padeZ	sreda
Cudey	kO:ru	padeZi	stansiya
InCeS	kaCeli	panedel'nik	stsena
Respublika)	kabinet	papiros	student
Sa: = bile	kambe:t	partiya	tabletka
Sartila:	kanikul	pi:va	taksi
Sina	karZi	piva	telo
Sveytsar	kartina	piyavka	teoriya
Zaket	kinZal	planeta	tetrad'
adres	kino	plotina	turist
aeroport	klimat	poema	turizm
akki	konfeta	poezd	uley
ami	korzina	poeziya	univermag
anglile:r	kreslo	politika	universitet
apteka	krolik	pomidor	viSnya
arenda	kude:	pressa	vinograd
argi:r	kvartira	problema	vintovka
armiya	kvitantsiya	professiya	voskresenye
asbest	lenta	provintsiya	vtornik
basseyn	limon	pulemyot	xONnu
buterbrod	maSina	pyatnitsa	xOlCok
dekabr	magnito(f)on	rabbit	xUlumzUrU:SkUn
dialekt	matematika	radio	xure:
e\'cadi	material	redaktorla:r	ok!kUske!kUske!
ekran	menyu	reklama	tir!baStaktanIp
ekzamen	metall	rekomentatsiya	boris
epos	metro	respublika	boris
fabrika	mifologiya	restoran	er!sallp
familiya	milyon	revolyutsiya	men!kIrIp
ferma	minut(a)	rifma	epCoksunup
fevral'	muzey	rolik	SayZe
garCitsa	myaCik	sekta	sovet

sovet	masterstvonu	ekonomiktig	dialekt
armiyaniN	Cexov	nacional	dialekt
revolyutsiyaniN	teatrI	dialektilerge	basseyneri
partiya	CexovtuN	dialektilerniN	dialektilerniN
partiyaniN	teatrni	dialektiniN	oCerkke
arxivinden	teatrniN	dialektizinge	dialektilerniN
aNmeN	scenazInga	dialektilerniN	dialektide
da:nCe	CexovtuN	dialektiziniN	dialektide
aziya	viSnyovIy	dialektilerniN	dialektide
geografitg	Cexov	dialektizmner	dialektide
aziyaga	gorkiy	dialektiniN	dialektide
evropaga	Cexov	dialektiziniN	dialektilderde
amerikaga	gorkiyeniN	dIlCe	dialektide
natsiya	Cexov	dialektilerniN	dialektide
teoriyalIg	maksim	dekabr	dialektide
materialdarga	gorkiyeni	dialekt	dialektide
Cexov	policiyaniN	dialektilerni	dialektide
Cexov	proletariy	dialekt	dialektilderde
gimnaziyani	revolyusCu	dialektiniN	dialektide
universitediniN	gorkiyeni	dialektige	dialektide
Cexov	nikolayga	dialektilerniN	dialektide
universiteti	rezolyuciya	dialektiler	dialektide
CexovtuN	vladimir	dialektileriniN	dialektilderde
Cexov	Cexov	dialektiler	dialektide
Cexov	CexovtuN	dialektileriniN	dialektileriN
povest	iyul	dialektilerden	sxemazI
garSin	Cexov	dialekt	dialekt
rossiyada	Cexov	dialekt	fonetikazI
Cexov	CexovtuN	dialekt	dialektige
Cexov	dialektilerniN	dialekt	dialektiden
CexovtuN	dialektilerniN	dialekt	nacional
Cexov	feodaldIg	dialekt	dialektiler
versta	administratsiya	dialekt	dIlCe
Cexov	dialektilerniN	dialekt	dialektilerden
Cexov	nacional	dialekt	dialekt
Cexovka	revolyutsiya	dialekt	dialektilerden
CexovtuN	politiktig	fonetikazI	daSce
Cexov	social	dialekt	daSce



**Appendix B:**  
**Igbo Disharmony Log**

nwetela	gwUworo	naekwuwapUta	mmetUta
naegosi	naeduga	naeme	naenyekarI
na:hUchasi	zopUta	emehiekwa”	gaetinye
naeduga	esiteghikwa	naeduhie	jizOs
emehiela	efesOs	egwU	onyenzOpUta
erUkwa	jizOs	matiu	efesOs
eme:la	onyenzOpUta	gaesi	ndahie
naeme	OmUma:tU/nlereanya	naenwe	jizOs
na:dIgide	cheta	pita	jizOs
na:dIgide	maobU	na:tachiri	naekwere/kwere
jizOs	emehiela	ntachiobi	jizOs
pita	jizOs	nchegharI”	nwurU
na:dIgide	kwesIrI	nkera	murU
naeme	ntUkwasIobi	naekwuwapUta	uwa”
ihUnanya	amen”	jizOs	akwUli
ndImmehie	kraist	jizOs	jizOs
jizOs	natawo	naewere	onyenzOpUta
kraist	gaesi	na:kUziri	jizOs
naewere	natawo	jizOs	inye/iwepUta
kwesIrI	gaesi	naewere	emehiela
naesote	ndIozi	jizOs	jizOs
naegosipUta	naekwupUta	wepUtara	kwesIrI
olileanya	naesite	naekwuwapUta	ntUkwasIobi
jizOs	kwusa:	banyere	amen”
kraist	mkpUrUokwu	obI	kraist
pita	ihicha	banyere	jizOs
ga:gbanwerIrI	ItōpU/IhapU	obI	onyenzOpUta
echegara	Ikagbu	banyere	jizOs
ndIozi	Ugwō	jizOs	onyenzOpUta
na:tUkwasi	weghachite	naekwuwapUta	natawo
kaonye	mmekOrIta	kOrint	jizOs
gaewere	na:gbanyeghi	jizOs	onyenzOpUta
kwupUta	ha/o	nkera	gaeburIrI
jizOs	merela	jizOs	“jizOs
gaewere	eme:la	inweta	“onyenzOpUta”
ga:zOputa	eklisiastis	naenwe	jizOs

jizOs	kristian	imekata	naekwerekwa
onyenkUzi	jizOs	nwekwaghI	jizOs
kwazi	kraist	agbanye	na:blakute
onyeamUma	mmekOrIta	sUwi:ch	jizOs
banyere	jizOs	jizOs	afOojuju
jizOs	onyenzOpUta	jizOs	jizOs
jizOs	ntUkwasIobi	kwuwapUtara	jizOs
IkUziri	nimegasI	naeso	naeduba
jizOs	ufOdU	ejegharI	naenye
onyenwenU/dinwenU	jizOs	gaenwe	afOojuju
jizOs	onyenzOpUta	naemekata	jizOs
onyenzOpUta	naenye	naemetUta	onyeOzUzUatUrU
onyenzOpUta	naeme	kpOchibidoro	jizOs
eme:la	nke/maka	abanye	jizOs
kwesIrI	jizOs	naenye	kewapUrU
kwesIrI	onyenzOpUta	afOojuju	emehiela
ekwesI	jizOs	jizOs	kewapUrU
na:dIgide	onyenzOpUta	jizOs	eklisiastis
onyenzOpUta	naeso	kwuwapUtara	naenwe
jizOs	cheta	anaeleda	mmetUta
jizOs	jizOs	mmekOrIta	naenwe
jizOs	gwUwororI:	gheretOkwa	mmekOrIta
jizOs	emehiela	naeji	kewapUrU
jizOs	jizOs	jizOs	mmekOrIta
gospitUtara	kwesIrI	jizOs	gaeme
ikwU	naetinye	onyeOzUzUatUrU	kewapU
jizOs	ntUkwasIobi	onyeOzUzUatUrU	jizOs
onyenzOpUta	jizOs	onyeOzUzUatUrU	jizOs
jizOs	onyenzOpUta	naechegharIba	naewere
onyenzOpUta	amen”	gaeme	naesote
jizOs	kraist	gafe:	jizOs
onyenzOpUta	emekata	gaere	naegosipUta
naelegara	metala	naemebe	jizOs
kristian	jizOs	ekwetacha	jizOs
imegasI	jizOs	jizOs	jizOs
naeme	naenye	jizOs	gaenwetezi
emegasI	na:blakute	kwuwapUtara	ga:gbanyegasI
kristian	naekwerekwa	naekwere	gaenwe

banye	naegosipUta	uwa	nkuwuwapUta
naenye	akOnauche	dirI	uwa
afOojuju	naeso	nelU	banyere
kachasi	jizOs	nleghara	nwegasIrI
onyeOzUzUatUrU	jizOs	mmadu	Oganihu
gaenwe	kristian	nwegasirI	nkuwuwapUta
kpOlitere	“kristian”	eme:la	ga:gbasi
na:nOnyere	ihuOma	gaenwejupUta	Ikwalite
jizOs	kristian	nlihapU	sitekwa
kraist	kristian	nokwa	debecha:
kristian	na:gachite	chOghi	Ikpasso
ōkOwa	naenyekwa	ikwanye	nwanne
dikshOnarI	kristian	ochIchI	nwanne
kristian	na:kUzi	ikpeazU	wepUtara
naegosipUta	anyi	gaechedo	nkuwuwapUta
gbadoro	naeme”	nwegasirI	banyere
nkUzi	kristian	Ikwalite	ozOkwa
jizOs”	mUworo	Oganihu	gbadoro
mmalite	pita	mmekOrIta	nlekOta
dikshOnarI	ntUkwasIobi	jikOtara	mmekOrIta
nIkUzi	jizOs	nations	nagbanyeghi
akwukwO	efesOs	kwupUtakwara	nlekOta
kristian	esiteghikwa	nwegasIrI	gaejigide
“kristian”	kristian	nwe:kwa	nOnOdu
testament	ntUkwasIobi	Ikwalite	odibO
ndIozi	jizOs	Oganihu	ga:machi
pita	kristian	eweputala	ga:kwagide
naeso	kristian	inweta	nazewetu
jizOs	ezinaUIO	njikO	nkwanyere
kristian”	jizOs	nations	tozukwa
ndIozi	kraist	nkwalite	inweta
naeme	nkuwuwaputa	nwegasIrI	tozukwara
“kristian”	uwa	banyere	nkuwuwapUta
mkipurUokwu	banyere	imejupUta	megidekwa
naeso	mmadu	nkwekOrIta	ikwanye
“kristian”	nwegasiri	general	ikpa
etufuola	mmalite	assembly	gaenwe
naeji	nghota	na:kwado	nkawado

nyegasIri	naeme	nwegasiri	enweghi
ga:kwagide	gaenyekwa	ga:kwalite	naedozi
mkpOchi	ItUliri	adimnamma	nketa
nwegasIri	inweta	kwalitekwa	njirimara
ida	tozukwa	nations	elektrOnik
karIri	inweta	naeme	njirimara
naenye	enyemaka	ozUzU	plastik
ga:kwagide	nhazi	gaenye	nketa
OcOghi	omenala	iketa	nketa
nzikOrIta	ɓula	nOganihu	nketa
bukwa	ga:ru	sayensi	nchekwa
nwe:kwa	inO	mmepUta	state
ilOghachi	inweta	sayensi	nketa
nwetakwa	enweghiOrU	nweta	nkwaru
gasi	gaeme	nbazi	nketa
mcpagbu	gbakwunyere	jikOtara	genweta
akpOku	Imalite	gaekwe	agbakwasi
nations	ibanye	gosipUtara	ezinUlo
gaesi	gaesokwa	nkuwuwaputa	nketa
nuzO	nlekOta	gaesi	gaesi
naedoghi	enyemaka	ngosipUta	gaenweta
Igbanwe	nlekOta	nwegasIri	nketa
Imalite	Ozo	gaenwe	saiba
nkwekOrIta	na:pUghi	mmachi	enweta
inwekOrIta	enyemaka	nweta	kpOmkwem
gaesi	nalUmdi	nwegasIri	kedU
naedoghi	mUghi	nweta	gaesi
gbakwunyere	gaenwe	ucheya	nketa
Igbanwe	ɓula	egosipUta	gaeme
igosiwapUta	inweta	gaesi	nketa
nkwupUta	mmalite	nations	ndebe
gbakwunyere	manyenye	nkuwuwapUta	nweta
IkwUdo	OzUzu	ga:kowa	ndebe
nagbanyeghi	gaerukwa	nketa	ndebe
da:manyenye	OzUzu	program	njirimara
nnOchite	gaeji	nketa	weghachiri
mmanyenye	nkwalite	naenyere	nketa
gaegosipUta	kwadosi	Ohaneze	ndebe

nwekwara	nwekwara	jizOs	naeduhie
nkea	kwadoro	kraist	egwU
saiba	inweta	naewere	matiu
nwale	nketa	kwesIrI	gaesi
nwale	nketa	naesote	naenwe
metutara	nwetaghI	naegosipUta	pita
nochita	ezinUIO	jizOs	ntachiobi
nketa	amerIka	kraist	nchegharI”
nketa	kwesIrI	pita	nkera
nlekOta	inweta	ndIozi	naekwuwapUta
enyemaka	nweta	kaonye	jizOs
akwadoro	nketa	gaewere	jizOs
nketa	amerIka	kwupUta	naewere
gaeme	nOdInihu	jizOs	jizOs
atUlecha:	kedU	gaewere	naewere
ndeba	ezinUIO	gwUworo	jizOs
nketa	gaesi	naeduga	wepUtara
gaenweta	Odinma	zopUta	naekwuwapUta
kedU	ezinUIO	jizOs	banyere
nketa	nketa	cheta	obI
enyemaka	ezinUIO	maobU	banyere
nketa	ichena	jizOs	obI
kpOo	oroma	kwesIrI	banyere
-enwetarIrI	osikapa	amen”	jizOs
njirimara	makroni	kraist	naekwuwapUta
amerIka	crackers	gaesi	kOrint
nketa	nenwechaghI	gaesi	jizOs
amanye	emegharI	ndIozi	nkera
ezinUIO		naekwupUta	jizOs
nketa	naegosi	naesite	naenwe
njirimara	naeduga	kwusa:	mmetUta
gaedeba	erUkwa	ItopU/IhapU	naenyekarI
nweta	naeme	Ugwo	gaetinye
enweta	jizOs	weghachite	jizOs
nketa	pita	mmekOrIta	ndahie
enweghI	naeme	ha/o	jizOs
eweta	ihUnanya	naekwuwapUta	jizOs
gaenwe	ndImmehie	naeme	naekwere/kwere

jizOs	mmekOrIta	mmekOrIta	naegosipUta
nwurU	jizOs	naeji	jizOs
murU	ufOdU	jizOs	jizOs
uwa”	jizOs	jizOs	jizOs
jizOs	naenye	naechegharIba	gaenwetezi
jizOs	naeme	gaeme	gaenwe
jizOs	nke/maka	gafe:	banye
kwesIrI	jizOs	gaere	naenye
amen”	jizOs	naemebe	gaenwe
kraist	naeso	jizOs	kpOlitere
jizOs	cheta	jizOs	jizOs
jizOs	jizOs	kwuwapUtara	kraist
jizOs	gwUwororI:	naekwere	ōkOwa
gaeburIrI	jizOs	naekwerekwa	dikshOnarI
“jizOs	kwesIrI	jizOs	naegosipUta
jizOs	naetinye	jizOs	gbadoro
jizOs	jizOs	jizOs	nkUzi
kwazi	amen”	jizOs	jizOs”
banyere	kraist	naeduba	mmalite
jizOs	metala	naenye	dikshOnarI
jizOs	jizOs	jizOs	akwukwO
jizOs	jizOs	jizOs	testament
jizOs	naenye	jizOs	ndIlozi
kwesIrI	naekwerekwa	kewapUrU	pita
kwesIrI	nwekwaghI	kewapUrU	naeso
jizOs	sUwi:ch	naenwe	jizOs
jizOs	jizOs	mmetUta	ndIlozi
jizOs	jizOs	naenwe	naeme
jizOs	kwuwapUtara	mmekOrIta	mkipurUokwu
jizOs	naeso	kewapUrU	naeso
ikwU	gaenwe	mmekOrIta	naeji
jizOs	naemekata	gaeme	naegosipUta
jizOs	naemetUta	kewapU	naeso
jizOs	kpOchibidoro	jizOs	jizOs
naelegara	naenye	jizOs	jizOs
naeme	jizOs	naewere	naenyekwa
jizOs	jizOs	naesote	anyi
kraist	kwuwapUtara	jizOs	naeme”

mUworo	assembly	nkwekOrIta	sayensi
pita	uwa	gaesi	mmepUta
jizOs	banyere	naedoghi	sayensi
jizOs	nwegasIrI	gbakwunyere	nweta
jizOs	nkwuwapUta	nkwapUta	nbazi
kraist	nwanne	gbakwunyere	jikOtara
uwa	nwanne	nnOchite	gaekwe
banyere	wepUtara	mmanye	gaesi
mmadu	nkwuwapUta	gaegosipUta	nwegasIrI
nwegasiri	banyere	naeme	gaenwe
mmalite	ozOkwa	gaenyekwa	mmachi
nghota	gbadoro	nhazi	nweta
uwa	nlekOta	hula	nwegasIrI
dirI	mmekOrIta	ga:ru	nweta
nelU	nlekOta	inO	gaesi
nleghara	gaejigide	gaeme	nations
mmadu	nazewetu	gbakwunyere	ga:kowa
nwegasiri	nkwanyere	ibanye	nketa
gaenwejupUta	nkwuwapUta	gaesokwa	program
nlihapU	ikwanye	nlekOta	nketa
nokwa	ikpa	nlekOta	naenyere
chOghi	gaenwe	Ozo	naedози
ikwanye	nyegasIrI	mUghi	nketa
ochIchi	mkpOchi	gaenwe	plastik
gaechedo	nwegasIrI	hula	nketa
nwegasiri	ida	mmalite	nketa
mmekOrIta	naenye	manye	nketa
jikOtara	nzikOrIta	gaerukwa	nchekwa
nations	bukwa	gaeji	state
kwupUtakwara	nwe:kwa	nkwalite	nketa
nwegasIrI	ilOghachi	kwadosi	nkwaru
nwe:kwa	nwetakwa	nwegasiri	nketa
njikO	gasi	adimnamma	nketa
nations	mcpagbu	kwalitekwa	gaesi
nkwalite	nations	nations	gaenweta
nwegasIrI	gaesi	naeme	nketa
banyere	nuzO	ozUzU	saiba
nkwekOrIta	naedoghi	gaenye	kpOmkwem

kedU	nkea	nketa	nketa
gaesi	saiba	nketa	nwetaghI
nketa	nwale	kpOo	amerIka
gaeme	nwale	amerIka	kwesIrI
nketa	nketa	nketa	nweta
ndeba	nketa	nketa	nketa
nweta	nlekOta	gaedeba	amerIka
ndeba	nketa	nweta	kedU
ndeba	gaeme	nketa	gaesi
weghachiri	ndeba	gaenwe	Odinma
nketa	nketa	nwekwara	nketa
ndeba	gaenweta	kwadoro	makroni
nwekwara	kedU	nketa	crackers

### Appendix C: Tuvan VHC Screenshot



## Results

10-19-2010 08:47pm  
Reducing diphthongs  
Long vowels enabled  
Vowels: i U e O I u a o  
Front vowels: i U e O  
Back vowels: I u a o  
Input file: Tuvan\_all.txt

Average number of syllables per word: 2.16.

Per polysyllabic word: 2.58.

---

Harmony threshold = 37.51%

Entire word harmony:

7390 out of 7763 words were harmonic.  
(95.20%)

373 out of 7763 words were disharmonic.  
(4.80%)

---

**Harmony index: 57.68%**

---

Harmony threshold for first two syllables =  
52.81%

Harmony in the first two syllables:

7533 out of 7763 words were harmonic.  
(97.04%)

230 out of 7763 words were disharmonic.  
(2.96%)

---

Vowel class distribution:

neutral: 0.00%

Front: 38.01%

Back: 61.99%

## Output files

The following files contain all of the information used in the VHC program's calculations. You can right-click these files and save them to your hard drive if you would like:

[Harmony log](#)

[Frequency log](#)

[Disharmony log](#)

## Results

10-19-2010 07:44pm  
Reducing diphthongs  
Long vowels enabled  
Vowels: i u e o I U a O  
Tense vowels: i u e o  
Lax vowels: I U a O  
Input file: Igbo\_data.txt

Average number of syllables per word: 1.85.

Per polysyllabic word: 2.44.

---

Harmony threshold = 37.05%

Entire word harmony:

4402 out of 5252 words were harmonic.

(83.82%)

850 out of 5252 words were disharmonic.

(16.18%)

---

**Harmony index: 46.76%**

---

Harmony threshold for first two syllables =

50.29%

Harmony in the first two syllables:

4712 out of 5252 words were harmonic.

(89.72%)

540 out of 5252 words were disharmonic.

(10.28%)

---

Vowel class distribution:

neutral: 0.00%

Tense: 44.00%

Lax: 56.00%

## Output files

The following files contain all of the information used in the VHC program's calculations. You can right-click these files and save them to your hard drive if you would like:

[Harmony log](#)

[Frequency log](#)

[Disharmony log](#)