Abstract
This study explores the upper bound of possible segments in the world’s languages and the features that define it. I limit the discussion to vowels in two databases of phoneme inventories: UPSID and P-Base. Common problems in using such databases are considered and a method to avoid them is introduced, where the notion of contrast plays a central role. The results show no compelling evidence for any language to require a three-way contrast in any feature dimension. In addition, while [high] and [low] are commonly used to define three degrees of vowel height, one is found to be sufficient, giving two degrees of height. I conclude that there are far fewer possible vowels, and far fewer features needed, than often assumed.

1. Introduction
The goal of this study is to explore whether we can define the set of possible segments, or consonants and vowels, in the world’s languages. I shall follow a common assumption, called ‘feature bounding’ by Clements (2009), that the set of possible segments is delimited by a set of features. For example, if there are N features, each having a binary value and able to combine freely with other features, there are 2^N possible segments. I shall not discuss all possible segments though. Instead, I shall focus on basic vowels. I begin with some preliminary questions: What are basic vowels? What are features and how are they determined? Do we have adequate data for the task? Can we compare sounds and features across languages?

1.1. Basic vowels
Basic vowels are those that involve lip rounding, the backness of the tongue, the height of the tongue, and the tongue root gesture. They roughly correspond to those in in an IPA chart, without diacritic marks for nasalization, breathiness, creakiness, etc.

Two questions require some discussion though. First, how is vowel length represented? There are three views. First, a long vowel is made of two short vowels. Second, a long vowel is distinguished from a short vowel by the feature [+long]. Third, long and short vowels differ in timing slots: a long vowel is linked to two timing slots and a short vowel is linked to one. I shall argue for the third position.

The second question concerns diphthongs and triphthongs. One approach treats them as single vowels. The other treats them as combinations of two or three vowels. In the first approach, Standard Chinese has twenty-one vowels (Lee and Zee 2003). In the second approach, Standard Chinese has five vowels (Duanmu 2007). The first approach makes little reference to syllable structure. For example, in Standard Chinese, [iau] rimes with [au]. This means that [iau] can be decomposed into [i] plus [au]. In addition, a short vowel can be followed by a consonant but a diphthong cannot. This suggests that a diphthong is equivalent to two sounds. In general, we can achieve a simpler analysis of syllable structure and a smaller inventory of vowels, if diphthongs and triphthongs are treated as clusters.

It is an open question whether diphthongs (and triphthongs) can be treated as two (or three) vowels in all languages. For example, some English speakers of New York City pronounce the vowel in *bath* and *cab* as [æʰ] (Cairns, p.c. 2013), which has been called a ‘short
diphthong’. However, it has been noted that the New York [æ] undergoes ‘tensing’ in such an environment (Benua 1995). If [æ] is long, it can be treated as a regular diphthong.

1.2. Features
Features serve two purposes (Halle 1962). The first is to distinguish sounds that are contrastive. The second is to define natural classes of sounds.

For the first purpose, a feature represents a minimal contrast between two sounds. A contrast is a difference between two sounds that can distinguish words in a language. Consider the examples in (1).

(1) Contrast in English

Minimal [s] vs. [z]  
sip vs. zip
Non-minimal [s] vs. [v]  
set vs. vet

It can be shown that the difference between [s] and [z] is minimal, commonly known as [voice], in the sense that [voice] cannot be further divided into two (or more) components each of which can itself be contrastive. It can also be shown that the difference between [s] and [v] is not minimal, in the sense it can be divided into two components, [voice] and ‘place,’ each of which can be contrastive by itself.

For the second purpose, there is an assumption that every natural class involves at least one feature value that is shared by all its members. For example, in English, the set of sounds that precedes the plural suffix [s] is a natural class and share the feature [-voice].

Ideally, features obtained from contrasts are the same as those obtained from natural classes. However, while linguists agree on what a contrast is, they do not always agree on what constitutes a phonological pattern or whether such a pattern always involves a natural class (see Mielke 2008). In this study, therefore, I focus on contrast only.

1.3. Adequacy of data
I shall use two of databases: UPSID (Maddieson and Precoda 1990) and P-Base (Mielke 2004-2007). UPSID contains 451 phoneme inventories and P-Base contains 628. Compared with the total number of languages in the world today, estimated to be at 6,000 (Moseley 2010), the databases may seem small. It is, therefore, natural to ask whether they adequate.

UPSID was compiled by selecting one language from each typological group. It is a reasonable representation of the world’s languages therefore. P-Base was compiled by collecting all inventories on which there is a published grammar book at two large university libraries (Ohio State University and Michigan State University). It is, therefore, also an objective representation of what we know.

Some linguists are optimistic that we already know enough. For example, Ladefoged and Maddieson (1996: 1-2) offer an upbeat statement below:

*We believe that enough is now known to attempt a description of the sounds in all the languages of the world... The 'global village' effect means that few societies remain outside the scope of scholarly scrutiny. In all probability there will be a sharp decrease in the rate at which previously unknown sounds are drawn to the attention of phoneticians.*
Besides the issue of coverage, other questions have been raised (Simpson 1999; Vaux 2009). For example, there are different treatments of diphthongs. Similarly, different analyses may choose different symbols to represent certain phonemes. For example, in P-base, Spanish has [b d g], but in UPSID they are given as [β δ γ]. A further issue is typographic convenience, as noted by Ruhlen (1976). For example, the IPA symbol [a] is supposed to be a low front vowel and [a] a low back vowel, but when a language does not have both, [a] is often used for [a]. Such issues pose problems if we are interested in the frequencies or the markedness of sounds (Clements 2009), but not if we are interested in contrasts among different sounds. For example, whether the Spanish sounds are [b d g] or [β δ γ], we may need to distinguish all of them, if some languages have all of them. Similarly, whether a low vowel in a given language is [a] or [a], as long as some language contrasts [a] and [a], we can capture the distinction. Moreover, our method will identify all inventories that appear to contain unusual contrasts, each of which will then be examined manually.

1.4. Cross-language comparison
A more serious question is whether sounds and features can be compared across languages. For example, both German and Norwegian use the same vowel symbols [i y e ø], yet the German vowels are systematically higher (Disner 1983: 67). Such small but systematic differences between languages are quite common. What is the reason to say that the Norwegian vowels are the same as those in German, beyond the fact that people happened to have chosen the same vowel symbols? Should such differences be distinguished at all?

Consider another common problem, illustrated with the backness of the tongue in (2), where A-D are four vowels in two hypothetical languages L1 and L2.

(2) Backness of four vowels A-D in two hypothetical languages L1 and L2

```
<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>L2</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>
```

The point of interest is the analysis of C. In L1, A is front and B back. If we consider L2 alone, we may call C front and D back. Can we then say then that A is the same vowel as C? Phonetically, they are different. However, if no language has more than two degrees of backness, we can consider A and C to be the same, both being front. If, however, we find a language that has three degrees of backness, we may need to reconsider whether C is front or central.

Given such issues, at least three views have been offered. According to the first (e.g. Ladefoged 1972; Disner 1983; Port and Leary 2005), each language is different and should be analyzed on its own. It makes little sense to identify sounds in one language with those in another. For example, in L1 of (2), we can call A front and B back, and in L2 we can call C front and D back, but it makes little sense to identify A with C, because ‘front’ in L1 does not mean the same as ‘front’ in L2. This approach fails to address the maximal number of contrasts in each phonetic dimension. In addition, this approach fails to appreciate the possibility that a mapping
relation can hold between sounds that are not phonetically identical, such as A and C in above, especially if no language has more than two degrees of contrast in backness, to be seen below.

According to the second view, sounds in different languages can be equated to each other, if we have a universal feature system (e.g. Chomsky and Halle 1968, for whom features are ‘substantive universals’). However, it remains to be shown how such a feature system is to be discerned from available inventory databases.

According to the third view, features can be derived from physical landmarks in the vocal tract (the ‘quantal theory’ of Stevens 1972). Therefore, at least some features can be identified across languages. It is unclear though how many degrees of contrast such a theory predicts for each phonetic dimension and how well the predictions fare against available databases.

The goal of this study is to gather empirical evidence on the maximal number of contrast in each phonetic dimension. Once that is known, we shall have some idea of the maximal number of possible segments. Therefore, the present study should be of interest to all parties in the theoretical debate.

2. Method
Our method follows three guidelines. I call them the Principle of Contrast, Known Feature First, and Maxima First. They are given in (3)-(5).

(3) The Principle of Contrast:
   a. If two sounds A and B can contrast in any language, they must be distinguished by at least one feature.
   b. If two sounds A and B never contrast in any language, they need not be distinguished by a feature.

(4) Known Feature First:
   Unless evidence requires otherwise, use known features or properties first before introducing a new feature.

(5) Maxima First:
   a. First, search through every language in order to determine the maximal number of contrasts in each phonetic dimension.
   b. Then, interpret the sounds in each language in terms of the maximal number of contrasts in each phonetic dimension.

2.1. Principle of Contrast
The Principle of Contrast is commonly used in the analysis of individual languages (e.g. International Phonetic Association 1999: 160). Our definition, however, extends it to cross-language comparisons. In particular, we can define allophones in terms of contrasts in other languages. This is shown in (6).

(6) Allophones: two sounds A and B (which have some phonetic similarity) are allophones of the same phoneme in a language if and only if
   a. A and B do not contrast in this language, and
   b. A and B contrast in another language.
Allophones are sounds that can potentially contrast in some language. If two sounds A and B never contrast in any language, they need not be distinguished as allophones in any language. For example, [m] (released) and [m] (unreleased) are sometimes listed as allophones in English. However, if they never contrast in any language, they need not be distinguished. Similarly, if [m nj] never contrast in any language, they need not be distinguished either.

There is clear evidence for the Principle of Contrast. For examples, consider (7), which shows eight vowels by two female speakers of American English, and (8), which shows three vowels (twenty tokens each) by one female speaker of American English.

(7)  [i i e æ a ɒ u] by two female speakers (solid line vs. broken line) of Midwestern American English, measured by the present author.

(8)  [i a u] (twenty tokens each) by one female speaker of American English, based on the narrow transcription of speaker s0101a from Columbus, Ohio, in the Buckeye Corpus (Pitt et al. 2007; measured by San Duanmu).
In (7), we see that some corresponding vowel pairs are quite different phonetically, yet their differences are ignored, since neither speaker considered the other to have any accent. In (8), we see that even for the same speaker, what are heard as [i u ø] by phonetic transcribers in fact vary a lot, which shows again that non-contrastive differences can be ignored.

Once we recognize the Principle of Contrast, we can use it for feature analysis. Consider Ao (Gowda 1991), which has [i u], and Apatani (Abraham 1985), which has [i i]. The IPA symbols suggest that there are three degrees of backness: front [i], central [i], and back [u], but neither language has a three-way contrast. If no other language has a three-way contrast, we can analyze the vowels in (9), where [i] in Apatani is reinterpreted as [u], and each language only has a two-way contrast in backness.

(9) Analysis of backness in two languages
Ao (Gowda 1991) i u
Apatani (Abraham 1985) i (i) ð u

2.2. Known Feature First
The purpose of Known Feature First is to minimize redundancy in representation. For example, consider the difference between [ə] and [ʌ] in English, shown in (10).

(10) Representing the difference between [ə] and [ʌ] in English
Feature difference [ə] central
[ʌ] back
Stress difference [ə] -stress
[ʌ] +stress

When [ə] and [ʌ] are distinguished, [ə] appears to be central and [ʌ] back. However, [ə] is an unstressed vowel and [ʌ] a stressed one. Since the distinction is already represented by stress, there is no need to represent it again by a feature (of backness).

Similarly, consider ‘advanced tongue root’ (ATR), ‘tense’, and ‘pharyngealized’. These features are similar in various ways. For example, among high vowels, tense correlates with advanced tongue root, or [+ATR], and lax correlates with retracted tongue root, or [-ATR], although the correlation is less obvious among low vowels (Halle and Stevens 1969). In addition, pharyngealized vowels (reported in !Xôô, Traill 1985) are made with retracted tongue root, or [-ATR]. Therefore, unless ATR, tense, or pharyngealized vowels contrast with each other in some language, we may not need all three features.

As a third case, consider vowel height (or factors that affect vowel height). Two options are shown in (11).

(11) Representing vowel height
1 feature 2 features
[i] high 1 +high, +ATR
[i] high 2 +high, -ATR
[e] high 3 -high, +ATR
[e] high 4 -high, -ATR
If we use one feature, we need a four-way contrast. If we use two features, we need a two-way contrast each. There is evidence that some languages need two features, such as Kinande (Kenstowicz 2009). Unless there is evidence otherwise, we can use two features for other languages, too.

2.3. Maxima First
As discussed above, without knowing the maximal number of possible contrasts in a phonetic dimension, it is difficult to compare sounds across languages. Maxima First offers a solution by setting up a system of reference for cross language comparisons.

Maxima First interacts with the Principle of Contrast, in that only contrastive differences in each feature dimension are represented. Maxima First also interacts with Known Feature First, in that when a language seems to show a larger than expected number of contrasts in a feature, we need to examine whether the contrasts can be represented with two (or more) known features, so that the number of contrasts in the original feature is reduced.

2.4. Procedure
Given the discussion above, we adopt the procedure in (12), where step (12d) involves the use of the Principle of Contrast and Known Feature First, to be illustrated below.

(12) Procedure of vowel analysis
   a. Extract a complete list of distinct vowel transcriptions.
   b. Divide the list into a set of basic vowels (those involving backness, height, rounding, and ATR) and those that are made of a basic vowel plus one or more additional features.
   c. Search through every language and extract inventories that seem to involve a controversial contrast (e.g. a three-way contrast in backness).
   d. Reexamine each extracted inventory and see if alternative analyses are available.

3. Result
I illustrate the process with data from UPSID. UPSID contains 269 distinct vowel transcriptions, totaling 3,833 tokens. They are divided into eight categories in (13). The category ‘laryngeal’ includes laryngeal (creaky) and breathy (murmur) vowels. The category ‘others’ includes voiceless, retroflex, and fricative vowels.
Vowels in UPSID.

<table>
<thead>
<tr>
<th>Category</th>
<th>Type</th>
<th>Token</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>45</td>
<td>2,699</td>
</tr>
<tr>
<td>Diphthong</td>
<td>89</td>
<td>201</td>
</tr>
<tr>
<td>Long</td>
<td>40</td>
<td>287</td>
</tr>
<tr>
<td>Nasalized</td>
<td>30</td>
<td>508</td>
</tr>
<tr>
<td>Laryngeal</td>
<td>24</td>
<td>68</td>
</tr>
<tr>
<td>Over-short</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>Pharyngeal</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>Others</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>All</td>
<td>269</td>
<td>3,833</td>
</tr>
</tbody>
</table>

Basic vowels refer to those that involve the backness of the tongue, the height of the tongue, lip rounding, and ATR (advanced tongue root, or tenseness) only. They are the most common vowels and cover 70% of all vowel tokens. Diphthongs and long vowels are composed of basic vowels and need not be discussed separately. Similarly, nasalized or laryngealized vowels need little elaboration, since features like nasalization, creakiness, and murmur can be added to a basic vowel. It can be shown, too, that the distinction between regular and ‘over-short’ vowels is similar to that between long and short vowels, since no language has a three-way contrast among over-short, regular, and long vowels. Therefore, this category is not discussed either. The category ‘pharyngeal’ is similar to ATR, where ‘pharyngealized’ corresponds to [-ATR] and ‘non-pharyngealized’ to [+ATR]; in addition, we found no language in which pharyngeal and ATR vowels contrast. Finally, the category ‘others’ involves voiceless vowels, retroflex vowels, and fricative vowels. No voiceless vowel is found to contrast with a regular vowel. Retroflex vowels can be represented with a coronal feature added to regular vowel features. Fricative vowels are found not to contrast with syllabic fricatives. Therefore, in what follows, I discuss basic vowels only.

UPSID assumes seven degrees of height, three degrees of backness, and two degrees of rounding. This gives 42 possible basic vowels, of which 38 are found, shown in (14).

<table>
<thead>
<tr>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>y</td>
</tr>
<tr>
<td>High (lower)</td>
<td>i</td>
<td>y</td>
</tr>
<tr>
<td>Mid (higher)</td>
<td>e</td>
<td>ø</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>ø</td>
</tr>
<tr>
<td>Mid (lower)</td>
<td>e</td>
<td>æ</td>
</tr>
<tr>
<td>Low (raised)</td>
<td>æ</td>
<td>æ</td>
</tr>
</tbody>
</table>
| Low   | a       | a    | a    | a    | a
Seven other basic vowels involve an additional diacritic and do not fit into (16). They are \([\text{\_ e \_ø \_y \_ə \_o} + \_+ i\}\], where \([\_ + -]\) indicate retracted, fronted, and velarized respectively. They are found not to contrast with \([\text{e e ø ø o i i}]\) and so require no further discussion.

Many phonologists assume just two degrees of backness and two degrees of height for mid vowels. Therefore, the controversial aspects of (14) are three degrees of backness and three degrees of height for mid vowels. A search was made for all triplets in these two features. There are twelve triplets for backness, shown in (15).

(15) Twelve contrastive triplets in backness
\[[i \ i u], [i \ i \ i], [\text{e ø y}], [\text{e ø y}], [\text{e ø y}], [\text{æ æ y}], [\text{a æ æ}]
[y u u], [y ø o], [ø ø o], [ø ø o], [œ œ ø] \]

A search through the inventories in UPSID yields three languages, with one triplet each. They are shown in (16).

(16) Search result for backness triplets

<table>
<thead>
<tr>
<th>Language</th>
<th>Triplet found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moro</td>
<td>[e ø y]</td>
</tr>
<tr>
<td>Nimboran</td>
<td>[i i u]</td>
</tr>
<tr>
<td>Woisika</td>
<td>[a æ æ]</td>
</tr>
</tbody>
</table>

Having gathered the exceptions, we examine each inventory to see if the case is valid. The vowel inventory of Moro is shown in (17), as given in the original source.

(17) Vowel inventory of Moro (Mr. and Mrs. Black 1971)

\[
\begin{array}{c|c}
\text{i} & \text{u} \\
\text{e} & \text{ə} & \text{y}, \text{o} \\
\text{a} & \\
\end{array}
\]

The triplet of interest is \([\text{e ø y}]\). However, according to the source, \([\text{ə}]\) occurs in unstressed positions only, whereas other vowels occur in stressed positions. This means that \([\text{ə}]\) is not a full phoneme and that Moro does not have a three-way contrast in backness.

Next we consider Woisika, whose vowel inventory is shown in (18), as given in UPSID, and (19), as given in the original source (Stokhof 1979)

(18) Vowel inventory of Woisika as given in UPSID

\[
\begin{array}{c|c}
\text{i, i} & \text{o, u} \\
\text{ɛ, e} & \text{ə, ø} \\
\text{a, a} & \text{æ, æ} \\
\end{array}
\]

(19) Vowel inventory of Woisika as given in the original source (Stokhof 1979)

\[
\begin{array}{c|c}
\text{i, i;} & \text{o, u;} \\
\text{ɛ, e;} & \text{ə, ø;} \\
\text{æ;} & \text{a;} \\
\end{array}
\]
The intended three-way contrast in backness is [a ą a], but as shown in the original source, there is a length difference between the central and back vowels, omitted in UPSID. If the difference between [a:] and [a] is represented by length, there is no need to represent it again by backness (central vs. back). Therefore, Woisika has no three-way contrast in backness.

Finally, consider Nimboran, whose vowels are [i i u e ę a], all of which are unrounded (Anceaux 1965: 9). The backness triplet is supposed to be [i i u]. However, according to the source (Anceaux 1965: 13-15), for some speakers [i] is ‘rather tense’ and ‘backed’, whereas [u] is slightly lowered. This means that [i u] could differ in tenseness or ATR, while both being back and high. Therefore, the Nimboran case does not seem compelling either.

Next, we consider three-way contrast in height among mid vowels. There are six such triplets, shown in (20). A search through UPSID yields two hits, shown in (21), both in the language Klao.

(20) Six contrastive triplets in height among mid vowels
[ɛ e ę], [ø o œ], [ə ə ɜ], [ɵ ɵ ɞ], [e o ɔ], [e o ɔ]

(21) Height triplets for mid vowels found in UPSID, both in Klao
[ɛ e ę], [ø o œ]

Klao has both oral vowels and nasal vowels. The oral vowels are shown in (22), as given in UPSID.

(22) Oral vowels in Klao, as given in UPSID

<table>
<thead>
<tr>
<th>High</th>
<th>Mid (higher)</th>
<th>Mid</th>
<th>Mid (lower)</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>ę</td>
<td>e</td>
<td>ę</td>
<td>a</td>
</tr>
<tr>
<td>u</td>
<td>o</td>
<td>o</td>
<td>ə</td>
<td>ɔ</td>
</tr>
</tbody>
</table>

However, in the source (Singler 1979: 63), the vowels are described differently, as shown in (23), where [e o] are [-ATR] (expanded pharynx) and [ę ə] are [-ATR].

(23) Oral vowels in Klao, as given in the original source (Singler 1979: 63)

<table>
<thead>
<tr>
<th>High</th>
<th>Mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>u</td>
<td>ø</td>
<td>ø</td>
</tr>
</tbody>
</table>

Of interest is the fact that [ę ə] are not mid but low vowels. Therefore, there is no three-way contrast in height among mid vowels.
In summary, we found no compelling case of three-way contrast in backness, or in the height of mid vowels. This is true of both UPSID and P-Base. Therefore, the maximal number of basic vowels is not forty-two but twenty, of which only nineteen are found, shown in (24), where there are two degrees of rounding and backness, and at most five degrees of height.

(24) Basic vowels in UPSID and P-Base

<table>
<thead>
<tr>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
</tr>
<tr>
<td>High (lower)</td>
<td>i</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
</tr>
<tr>
<td>Mid (lower)</td>
<td>ε</td>
</tr>
<tr>
<td>Low</td>
<td>æ</td>
</tr>
</tbody>
</table>

4. A close look at vowel height
According to Chomsky and Halle (1968), Kiparsky (1974), and others, there are five binary features for basic vowels, [back], [round], [high], [low], and [ATR]. The system yields twenty-four basic vowels, shown in (25).

(25) Basic vowels proposed by Chomsky and Halle (1968) and Kiparsky (1974)

<table>
<thead>
<tr>
<th>-round</th>
<th>+round</th>
<th>-round</th>
<th>+round</th>
<th>-round</th>
</tr>
</thead>
<tbody>
<tr>
<td>+high, -low</td>
<td>+ATR</td>
<td>i</td>
<td>y</td>
<td>u</td>
</tr>
<tr>
<td>-ATR</td>
<td>i</td>
<td>y</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>-high, -low</td>
<td>+ATR</td>
<td>e</td>
<td>ø</td>
<td>γ</td>
</tr>
<tr>
<td>-ATR</td>
<td>e</td>
<td>ø</td>
<td>γ</td>
<td>o</td>
</tr>
<tr>
<td>-high, +low</td>
<td>+ATR</td>
<td>æ</td>
<td>ð</td>
<td>Ł</td>
</tr>
<tr>
<td>-ATR</td>
<td>æ</td>
<td>ð</td>
<td>Ł</td>
<td>ð</td>
</tr>
</tbody>
</table>

The system is more parsimonious than most others, such as that of a standard IPA table (International Phonetic Association 1999) or that of UPSID. Still, the system exceeds what is needed to represent all vowel contrasts in UPSID and P-Base. If we use the same features, we can represent our results in (26), where some vowel symbols are slightly adapted and [ATR] is not used for low vowels.
(26) Basic vowels proposed by Chomsky and Halle (1968) and Kiparsky (1974)

<table>
<thead>
<tr>
<th></th>
<th>-back</th>
<th>+back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-round</td>
<td>+round</td>
</tr>
<tr>
<td>+high, -low</td>
<td>+ATR</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>-ATR</td>
<td>i</td>
</tr>
<tr>
<td>-high, -low</td>
<td>+ATR</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>-ATR</td>
<td>e</td>
</tr>
<tr>
<td>-high, +low</td>
<td>ə</td>
<td>ø</td>
</tr>
</tbody>
</table>

While (26) is simpler, the need for three degrees of height has not been demonstrated in our study. In addition, there are several other problems. First, there is a missing vowel in (26). Second, height seems to be the only phonetic dimension that has three degrees of contrast (even though it is represented with two features). Third, it remains unclear why [ATR] does not apply to low vowels. Fourth, while minimally contrastive pairs between high and mid vowels are easy to find, minimally contrastive pairs between mid and low vowels are quite rare (see English examples below). Finally, there are languages where vowels fall into just two groups of height, even when they seem to show three degrees of height phonetically. This is the case in Turkish, whose vowels are shown in (27), where [a] is low and central.

(27) Vowels in Turkish (Zimmer and Orgun 1992: 44).

However, with regard to vowel harmony, Turkish vowels fall into two degrees of height, shown in (28), where [a] is back and belongs to the same height category as phonetically mid vowels, both called ‘open’ (Lewis 1967).

(28) Feature analysis of Turkish vowels (Lewis 1967: 14)

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrounded</td>
<td>Rounded</td>
</tr>
<tr>
<td>Close</td>
<td>i</td>
<td>y</td>
</tr>
<tr>
<td>Open</td>
<td>e</td>
<td>ø</td>
</tr>
</tbody>
</table>

It is worth asking, therefore, whether fewer basic vowels are sufficient to account for the inventories in UPSID and P-Base. If we only assume two degrees of height, the number of basic vowels is sixteen. This is shown in (29).
Inventory of basic vowels in a two-height analysis

<table>
<thead>
<tr>
<th></th>
<th>[-back]</th>
<th>[+back]</th>
<th></th>
<th>[-round]</th>
<th>[+round]</th>
<th></th>
<th>[-round]</th>
<th>[+round]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+high]</td>
<td>[+ATR]</td>
<td>i</td>
<td>y</td>
<td>u</td>
<td>u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-ATR]</td>
<td></td>
<td>i</td>
<td>y</td>
<td>ũ</td>
<td>ũ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-high]</td>
<td>[+ATR]</td>
<td>e</td>
<td>œ</td>
<td>γ</td>
<td>o</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-ATR]</td>
<td></td>
<td>æ</td>
<td>æ</td>
<td>a</td>
<td>c</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The choice of the phonetic symbols is flexible, but it has little consequence for our discussion. Let us now consider whether a sixteen-vowel system is sufficient to account for all vowel inventories. We shall look at British English, German, Swedish, and !Xóõ. The first three are chosen because their patterns are well known and their vowel inventories are fairly large. !Xóõ is chosen because it has the largest vowel inventory in P-Base. We shall focus on whether (31) offers enough positions, rather than which position each vowel should go into.

British English has ten monophthongs. A possible analysis is shown in (30), following the transcription of Ladefoged (2001: 29). It is worth noting that there is phonological evidence for vowel length in English, both in syllable structure (Borowsky 1986) and in stress assignment (Halle and Vergnaud 1987; Hayes 1995).

Next we consider German, which has nineteen vowels (Wiese 1996; Kohler 1999, Fox 2005). They can be analyzed in (31), excluding [ə], which occurs in unstressed syllables only, and three diphthongs.

The two-height system has more than enough slots to accommodate English vowels. The vowels [æ, a] can share the same slot, because they differ in length. The analysis shows that [tense] does not always correspond to [ATR]. Instead, length is a better correlate of [tense].

Next we consider German, which has nineteen vowels (Wiese 1996; Kohler 1999, Fox 2005). They can be analyzed in (31), excluding [ə], which occurs in unstressed syllables only, and three diphthongs.

[ɛ ɛː] and [ə aː] show again that [ATR] is independent from length. In addition, as in English, there is phonological evidence for vowel length in German (Giegerich 1985).
Next we consider Swedish, which has seventeen vowels. They can be analyzed in (32), according to the transcription of Engstrand (1999: 141).

(32) Two-height analysis of Swedish vowels

<table>
<thead>
<tr>
<th>i:</th>
<th>y:</th>
<th>ŋ, u:</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>y</td>
<td>ŋ</td>
</tr>
<tr>
<td>e:</td>
<td>ø:</td>
<td>o, o:</td>
</tr>
<tr>
<td>e, e:</td>
<td>œ</td>
<td>a, a:</td>
</tr>
</tbody>
</table>

Some positions are again filled by two vowels each, which differ in length, and there is no need to distinguish them by another feature.

Finally, let us consider !Xóõ, which has forty-four vowels, the largest vowel inventory in P-Base. The vowels are shown in (33).

(33) Vowels in !Xóõ, the largest inventory in P-Base

| i | e | a | o | u | ē | ā | ō | ū | ĭ | ē | ā |
| q | ŭ | ã | õ | ũ | ı̃ | ẽ | ã | õ | ũ | ē | ā |
| ŏ | ŭ | ã | õ | ŭ | ĭ | ŏ | ŭ | ē | ŏ | ŭ | ū | ā | ū | ã | ũ |

Nasalization, murmur, and glottalization are non-basic features. In addition, there is no contrast between pharyngealization and ATR, in that pharyngeal vowels can be seen as [−ATR] and other vowels as [+ATR]. Thus, !Xóõ has just eight basic vowels, analyzed in (34).

(34) Two-height analysis of basic vowels in !Xóõ

<table>
<thead>
<tr>
<th>i</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>ũ</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>a</td>
</tr>
<tr>
<td>ã</td>
<td>õ</td>
</tr>
</tbody>
</table>

How many languages might have more than sixteen basic vowels? To find out, consider the sizes of all vowel inventories in P-Base, shown in (35).

(35) Vowel inventory sizes in P-Base

<table>
<thead>
<tr>
<th>Vowel inventory size</th>
<th>Raw</th>
<th>Basic vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 16</td>
<td>595</td>
<td>627</td>
</tr>
<tr>
<td>Larger than 16</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>628</td>
<td>628</td>
</tr>
</tbody>
</table>

Out of 628 inventories, just 33 appear to have more than sixteen vowels. If we exclude non-basic features (such as length, nasalization, murmur, etc.), just one language, Turkana, has more than sixteen vowels. The Turkana inventory is shown in (36).
In Turkana, there are just nine basic vowels, the rest being their voiceless counter-parts. The source, Dimmendaal (1983), notes that voiceless vowels lack stress. In a later analysis, Dimmendaal (1993: 131) no longer includes them in the phonemes inventory.

In summary, there is no compelling evidence for three degrees of vowel height, and sixteen basic vowels seem sufficient to account for all inventories in P-Base and UPSID.

5. Concluding remarks
The goal of this study is to determine the range of possible segments in the world’s languages, using UPSID and P-Base as our data. I have shown that, as far as basic vowels are concerned, only four binary features are needed: [back], [high], [round], and [ATR]. This system yields a total of sixteen basic vowels, shown in (29). In other words, there are far fewer possible vowels that commonly assumed, and far fewer features to distinguish them.

Our discussion has focused on contrast, with only occasional references to natural classes. With fewer features than before, our system shall be unable to account for certain natural classes. How then should the two systems, one based on contrast and one on natural classes, be reconciled?

It is worth noting that a contrast-based system is a minimal system that all phonologists would recognize. The question is whether we should expand the system in order to account for natural classes. The answer is not obvious, but I shall offer some speculations.

One possibility is that the contrast-based system should be expanded with additional features to account for natural classes. Such additional features could serve as potential contrasts, too, which happen not to be used in the languages we examined.

The other possibility is that natural classes not definable with contrast-based features need to be reexamined. This is a strong claim, but not entirely new. For example, in a large-scale study on natural class-based features, Mielke (2008) reports that 30% of what are thought to be natural classes are ‘unnatural’, in the sense that they cannot be defined by well-known features. Clearly, much work is needed, and it will be interesting to find out how to resolve the differences between evidence from contrast and that from natural classes.
References:


The Hague: Mouton.


