Phonology

A phoneme is a distinct speech sound in a language.

Lessons

Minimal Pairs
  Overview
  Example Problems
  Exercises

Phonemes and Allophones
  Definitions
  Explanation
  Summary

Complementary & Contrastive Distribution
  Definitions
  Explanation
  Summary

Environments
  Overview
  Step-by-Step Example
  Example Problems
  Exercises

Determining the Underlying Phoneme
  Overview
  Example Problems
  Exercises

Rules
  Overview
  Example Problems
  Reference: Rule-writing Conventions
  Exercises

Analyzing Data Sets
  Overview
  Example Problems
  Exercises

Phonological Processes
  Overview
  Example Problems
  Exercises
**Minimal Pairs**

*Objective:* To identify minimal pairs.

Minimal pairs are pairs of words with different meanings and exactly one sound difference. For example, “cat” and “bat” are minimal pairs because only the first sound is different ([k] vs [b]). However, “cat” and “flat” are not minimal pairs, because there are two sound differences: ([k] vs [f] and [l]).

The reason we look for minimal pairs is to identify a contrast between two sounds. If two different sounds, placed in the same exact environment, produce different words with different meanings, then those sounds really are different phonemes. See the “Phonemes vs. Allophones” lesson in this section for more information.

Remember, the restrictions on minimal pairs are:

- The two words have different meanings
- Only one sound is different
- The words have the same number of sounds
- The sound that is different is in the same place in both words

**Tips**

- If the words are in English, they will probably be given to you in standard orthography. Remember, DO NOT RELY ON SPELLING! Write the words in IPA. Then you just have to compare the symbols. For example, in standard orthography it looks like “rewind” and “resigned” must have more than one sound change, but if you write them in IPA (ræwənd, rəzænd) you will see that they are actually minimal pairs (Again, your dialect may vary.)
- If the words are in another language, you will be given the transcription and a gloss (what the word means). So, you don’t have to worry about trying to do a transcription into IPA.

**Example Problems**

1. Are the following pairs of English words minimal pairs?
   a. raven, craven
      IPA transcriptions: rævən, krevən
      These are not minimal pairs—they don’t have the same number of sounds. The idea is to compare one sound with another, but there is no sound to compare [k] to, because it corresponds to no sound in the word “raven.”

   b. cab, cash
      IPA transcriptions: kæb, kæʃ
      These are minimal pairs. You’re comparing [b] and [ʃ].
2. Consider this data set from another language, and determine whether the given pairs are minimal pairs or not.
   a. Tagalog:  
      kahon  “box”
      kaʔon  “to fetch”
      These two words have different meanings and one sound change (between “h” and “ʔ”).
      They are a minimal pair.
   
   b. Inuktitut:  
      ɪglu   (snow)house
      ɪglu   (snow)house
      These two words have the same meaning, so even though there is a sound change, they do not constitute a minimal pair.

Exercises

1. Are the following pairs of English words minimal pairs?
   a. law, jaw
   b. crime, time
   c. prison, prism

2. Consider this data set from Thai, and determine whether the given pairs are minimal pairs or not.
   a. ɓat (“sheet”), ɗat (“to go”)
   b. bryy (“extremely fast”), myy (“hand”)
   c. ɗat (“to go”), ɗʰat (“danger”)
Phonemes vs. Allophones

Objective: To distinguish phonemes from allophones.

Definitions
So far, we’ve been describing speech sounds. Now we’re going to distinguish between two types of speech sounds: phonemes and allophones.

Phoneme: A speech sound that is distinct from other sounds in the language. Changing a phoneme changes the meaning of a word.

Allophone: A speech sound that is a variation of some phoneme in a language. An allophone of a phoneme is a version of that phoneme which is always found in some particular environment.

Explanation
There is a lot of variation in the way that sounds are produced. For example, someone with a very high voice saying [k] creates quite a different sound wave than someone with a very low voice saying [k]. Yet we still recognize the two different sounds as being the same; they’re the same phoneme. There’s any number of individual or random variations that affect the quality of the sound, but which don’t affect us as listeners and perceivers of language. We’re wired to process phonemes, and string them into meaningful words in our minds, without noticing unimportant differences in how they sound from one speaker or moment to the next. What differences count as unimportant? Well, it depends on the language.

[s] and [f]: Phonemes in English, allophones in Korean
Consider the difference between [s] and [f]. (Quick, look at your IPA consonant or feature charts. What is the difference?) In English, we consider the distinction to be an important one: [s] and [f] are different phonemes. One way to tell is that minimal pairs like “soot” and “shoot” have different meanings. Take a word, change the [s] sound to a [f] sound, and you’ve got either a different word (“sop” becomes “shop”) or meaningless nonsense (“soap” becomes “shoap”).

So, [s] and [f] are definitely different phonemes in English, but that isn’t so in all languages. For example, in Korean, [s] and [f] are allophones of the same phoneme. Changing [s] to [f] does not change the meaning of the word; it just sounds a little strange.

[p] and [b*]: Allophones in English, phonemes in Thai
The different symbols in the IPA chart you have represent sounds that are different phonemes in English. Now we’ll describe a new set of sounds which are not in your IPA chart, but which you already know how to make if you are a native speaker of English—in fact, you use them all the time. These are the aspirated stops.

Consider the difference between words like “pot” and “spot”. Say them, and really pay attention to how they sound, and how they feel. Can you tell the difference between the [p] is “pot” and in “spot”? Try really overemphasizing them. If you’re still not sure, try holding a piece of paper in front of your mouth, or speaking at a lit candle. The paper, or the flame, should move when you say a word-initial [p] (“pa, pa, pa, pa”) and stay still when you say [p] after [s] (“spa,
spa, spa, spa”). (If it doesn’t, you may not have this difference with your particular dialect or accent. Don’t worry—someone in your class will have it, and you don’t need to have the difference to understand what it is.) The two kinds of [p] are really different sounds: one has an additional puff of air, which is called aspiration. Aspirated [p] like you find at the beginning of a word is written as [pʰ], whereas the unaspirated [p] like you find in “spot” is written just as regular old [p]. You can find aspiration with every voiceless stop in English (so, [t] and [k] also). The candle trick also works with “stop” and “top,” or “cab” and “scab,” for example.)

Now, in English, these are not different phonemes. If you really try, you can get aspiration on a voiceless stop after [s] (try saying [stʰap]) or produce an unaspirated word-initial voiceless stop (try saying [tap])—it sounds weird to native speakers, but it doesn’t change the meaning of the word.

There are some languages in which aspiration is a phonemic difference, though. For example, in the minimal pair exercise above, the Thai words pat (“to go”) and pʰat (“danger”) constitute a minimal pair which contrasts aspirated and unaspirated [p]. Since the meaning of the word changes depending on which [p] is used, aspiration is a phonemic difference in Thai.

Psycholinguistic Research
How can phonemic distinction different cross-linguistically? Aren’t sounds like [p] and [b] fundamentally different? In some ways, the distinctions we draw between different sounds are arbitrary. Take a pair like [r] and [l]. By slowing moving your tongue forward and back, you can switch continuously between the two sounds. They can be thought of not so much as two separate, concrete entities, but as ends of a continuum. Voiced/voiceless pairs are also like that. Suppose someone is making a bilabial stop followed by [a] (so, “pa” or “ba”). Because of the vowel, the speaker will have to start voicing at some point. Whether the listener hears a [p] or a [b] depends on the voice onset time. If voicing starts early, the listener will hear “ba”; if voicing starts late, they will hear “pa”. But if the voicing starts at some in-between time, they have to make a judgment call.

The extent to which various sounds seem distinct varies from language to language. It may sound like [r] and [l] are very different because we are used to them being different; however, people who speak languages that don’t treat [r] and [l] as different phonemes find that they sound very similar. In contrast, it may seem to native English speakers that [p] and [pʰ] sound very similar, but to speakers of Thai, they sound entirely different.

Summary
- Even though there may be systematically different ways of pronouncing some sound (that is, different allophones of some phoneme), speakers of a language recognize the difference between different phonemes and usually fail to notice the difference between different allophones.
- Whether a pair of sounds represents two different phonemes or allophone variants of the same phoneme depends on the language.
**Complementary and Contrastive Distribution**

*Objective:* Distinguish between contrastive and complementary distribution. Know which one indicates phonemes, and which indicates allophones.

**Definitions**

**Environment:** For these simple examples, the environment is simply the sounds before and after the ones we are looking at.

**Contrastive Distribution:** Sounds that are in contrastive distribution can be found in the same environment. The sounds *contrast* and therefore they are different phonemes.

**Complementary Distribution:** Sounds that are in complementary distribution are always found in different environments. The two sounds *complement* each other—that is, between them, they cover all possible environments. They are allophones of the same phoneme.

**Explanation**

We know that different phonemes can appear in the same surrounding word and create different meanings (minimal pairs). We also know that when there are different allophone variants of a phoneme, the environment determines which one appears (for example, in standard English, [pH] is always pronounced at the beginning of a word, never [p]). We can use this knowledge to determine whether the difference between a pair of sounds is phonemic or allophonic in a given language. If the two sounds can appear in the same environment (that is, with the same surrounding sounds), then they are in *contrastive distribution* and they are different phonemes. On the other hand, if the two sounds *always* appear in different environments—if there’s a systematic way to tell when to use one sound and when to use the other—then they’re in *complementary distribution* and they’re allophones of the same phoneme.

**Summary**

- Phonemes are found in *contrastive distribution* (same environment)
- Allophones are found in *complementary distribution* (different environments)
Environments

Objective: Make charts showing the environments to compare two sounds, and make a decision about complementary vs. contrastive distribution. If complementary, decide which is the underlying phoneme.

Overview
In order to determine if a pair of sounds represents different phonemes or different allophones, we need to know whether they are in contrastive or complementary distribution. And in order to determine that distribution, we need to carefully look at the environments in which the sounds are found. If the sounds can be found in the same environment—if both sounds could come before and after the same surrounding sounds—then they’re in contrastive distribution and they’re phonemes. But if we can find some systematic difference in the environments, then we’re looking at complementary distribution (allophones).

If there are minimal pairs that contrast the sounds you’re looking at, you know right away that the sounds are different phonemes. If there aren’t, though, you need to carefully look at the sounds before and after. One way to clarify the problem is to draw charts of the immediate environments of each sound.

Representing Environment
The environment of a sound means the sounds that come immediately before and after it. By convention, we write this as [preceding sound]_[following sound]. For example, the environment of [k] in the word [bækət] would be written s_ə.

Environment-Writing Tips:
- Word boundaries—the beginning or end of a word—are represented by the symbol #.
  The environment of the [k] in [lʊk] (the name “Luke”) would be written u_#.
- A diphthong counts as one sound, so the environment for [k] in [skəj] is s_əj.
- Notice that a single word can have the letter of interest multiple times. For example, [lʊk skəjwəkər] has three separate environments for [k]:
  u_# s_əj s_ə

Environment Charts: Step-by-Step
1. Make a list of the environments for each sound of interest.
   You will typically be comparing two sounds found in a data set. For each sound, go through the data set, writing down every environment where that sound is found. Keep the lists separate.

2. Compare each side of one list to the corresponding side of the other list, looking for overlap.
   Overlap means the same sounds appear in the same position (before or after the sound of interest) on both lists.
   If there is overlap on both sides, we can conclude the sounds are in contrastive distribution because they could, theoretically, appear in the exact same environment. There’s no difference between the environments in which one sound occurs vs. the other sound, so the sounds themselves contrast. In other words, they’re two different phonemes.
If there is no overlap on one or both of the sides (i.e. there is no overlap on either side OR either the right side or left side has no overlap, although the other may certainly have overlap), we can conclude the sounds are in complementary distribution, because they occur in systematically different environments. In other words, they are two allophone versions of the same underlying phoneme. Later sections will show how to tell which is the underlying phoneme and what rule governs the use of the allophone.

Example Problems

1. Consider this data from Tongan, a Polynesian language. Are [s] and [t] different phonemes, or allophones of the same phoneme?

<table>
<thead>
<tr>
<th>Tongan</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>tauhi</td>
<td>“to take care”</td>
</tr>
<tr>
<td>sisi</td>
<td>“garland”</td>
</tr>
<tr>
<td>motu</td>
<td>“island”</td>
</tr>
<tr>
<td>mosimosi</td>
<td>“to drizzle”</td>
</tr>
<tr>
<td>motomoto</td>
<td>“unripe”</td>
</tr>
<tr>
<td>fesi</td>
<td>“to break”</td>
</tr>
<tr>
<td>sino</td>
<td>“body”</td>
</tr>
<tr>
<td>totonu</td>
<td>“correct”</td>
</tr>
<tr>
<td>pasi</td>
<td>“to clap”</td>
</tr>
<tr>
<td>fata</td>
<td>“shelf”</td>
</tr>
<tr>
<td>movete</td>
<td>“to come apart”</td>
</tr>
<tr>
<td>misi</td>
<td>“to dream”</td>
</tr>
</tbody>
</table>

First off, we can check for minimal pairs, since that’s the easiest way to tell that the two sounds are phonemes. There are none, so we can’t conclude anything yet. We need to figure out if the sounds are in complementary or contrastive distribution.

Step 1. Make a list of environments for each sound of interest.

Our chart looks like this:

```
  s
#_i  t
 i_i  #_a
 o_i  o_o
 o_i  o_o
 e_i  #_o
 #_i  o_o
 a_i  a_a
 i_i  e_e
```

Note that since we care about variability, there’s really no point in listing all the repeated environments in each list. So we could write the list like this:

```
  s
#_i  t
 i_i  #_a
 o_i  o_u
 o_i  o_o
 e_i  #_o
 a_i  a_a
```

26
In the rest of the examples we present, we won’t write repeat environments. For the purposes of comparing the lists, it doesn’t matter whether you do or not. Just make sure, if you get rid of them, that you do it separately for each list—so if the same environment is in both lists, you definitely want to leave that.

**Step 2. Compare each side of one list to that side of the other list, looking for overlap.**

Let’s look at the lists we made in the previous step. First, we will compare the left-hand side of each environment. (We re-wrote the lists one side at a time for this purpose, but you can do this visually in a number of ways—covering up the side you’re not looking at, circling the side you are looking at, etc.)

**Left-hand side comparison**

<table>
<thead>
<tr>
<th>s</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>#</td>
</tr>
<tr>
<td>a</td>
<td>e</td>
</tr>
</tbody>
</table>

There’s definitely overlap between these two lists: we see [o], [e], [a] and word boundaries (#) on both lists.

What does this mean? Well, if we find overlap on the right-hand side as well, then the sounds are in contrastive distribution (and are phonemes). However, if there is no overlap on the right-hand side, then we can conclude that they are in complementary distribution, and it’s the sound on the right-hand side that causes the change from one allophone to the other. So let’s check out the right-hand side.

**Right-hand side comparison**

<table>
<thead>
<tr>
<th>s</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>a</td>
</tr>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>i</td>
<td>o</td>
</tr>
<tr>
<td>i</td>
<td>a</td>
</tr>
<tr>
<td>i</td>
<td>e</td>
</tr>
</tbody>
</table>

Absolutely no overlap. The list for [s] consists only of [i], and there are no [i]s on the list for [t]. Since we found no overlap on one side, we can conclude that the two sounds are in complementary distribution. They’re allophones of the same phoneme.
2. Using the same data set from the previous problem (reprinted below), are [m] and [t] different phonemes, or allophones of the same phoneme?

<table>
<thead>
<tr>
<th>Words</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>tauhi</td>
<td>“to take care”</td>
</tr>
<tr>
<td>sisi</td>
<td>“garland”</td>
</tr>
<tr>
<td>motu</td>
<td>“island”</td>
</tr>
<tr>
<td>mosimosi</td>
<td>“to drizzle”</td>
</tr>
<tr>
<td>motomoto</td>
<td>“unripe”</td>
</tr>
<tr>
<td>fesi</td>
<td>“to break”</td>
</tr>
<tr>
<td>sino</td>
<td>“body”</td>
</tr>
<tr>
<td>totonu</td>
<td>“correct”</td>
</tr>
<tr>
<td>pasi</td>
<td>“to clap”</td>
</tr>
<tr>
<td>fata</td>
<td>“shelf”</td>
</tr>
<tr>
<td>movete</td>
<td>“to come apart”</td>
</tr>
<tr>
<td>misi</td>
<td>“to dream”</td>
</tr>
</tbody>
</table>

First, we check for minimal pairs, since that’s the easiest way to tell that the two sounds are phonemes. There are none, so we can’t conclude anything yet. We need to figure out if the sounds are in complementary or contrastive distribution.

**Step 1. Make a list of environments for each sound of interest.**

Our chart looks like this:

<table>
<thead>
<tr>
<th>t</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_a</td>
<td>#_o</td>
</tr>
<tr>
<td>o_u</td>
<td>i_o</td>
</tr>
<tr>
<td>o_o</td>
<td>o_o</td>
</tr>
<tr>
<td>#_o</td>
<td>#_i</td>
</tr>
<tr>
<td>a_a</td>
<td></td>
</tr>
<tr>
<td>e_e</td>
<td></td>
</tr>
</tbody>
</table>

**Step 2. Compare each side of one list to that side of the other list, looking for overlap.**

First, we will compare the left-hand side of each environment.

**Left-hand side comparison**

<table>
<thead>
<tr>
<th>t</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>o</td>
<td>i</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>#</td>
<td>o</td>
</tr>
<tr>
<td>a</td>
<td>#</td>
</tr>
<tr>
<td>e</td>
<td></td>
</tr>
</tbody>
</table>

There’s overlap: [o] and word boundaries (#) occur on both lists.

What does this mean? Well, if we find overlap on the right-hand side as well, then the sounds are in contrastive distribution (and are phonemes). However, if there is no overlap on the right-hand side, then we can conclude that they are in complementary distribution, and it’s the sound on the right-hand side that causes the change from one allophone to the other. So let’s check out the right-hand side.
Right-hand side comparison

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>m</td>
</tr>
<tr>
<td>a</td>
<td>o</td>
</tr>
<tr>
<td>u</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>i</td>
</tr>
<tr>
<td>a</td>
<td>e</td>
</tr>
</tbody>
</table>

There’s overlap on this side, too—[o] occurs on both lists. Since we see overlap on both sides, the sounds are in contrastive distribution, and we can conclude that [t] and [m] are different phonemes in this language.

Exercises

1. Consider the following data from English (use the IPA transcriptions, not the standard orthography glosses.) According to this data, are [θ] and [ð] allophones of the same phoneme, or different phonemes?

<table>
<thead>
<tr>
<th>boθ</th>
<th>“both”</th>
</tr>
</thead>
<tbody>
<tr>
<td>doz</td>
<td>“those”</td>
</tr>
<tr>
<td>rædɔr</td>
<td>“rather”</td>
</tr>
<tr>
<td>θm</td>
<td>“thin”</td>
</tr>
<tr>
<td>fædɔrz</td>
<td>“fathers”</td>
</tr>
<tr>
<td>θɪŋk</td>
<td>“think”</td>
</tr>
<tr>
<td>ðeɪ</td>
<td>“they”</td>
</tr>
<tr>
<td>loð</td>
<td>“loathe”</td>
</tr>
<tr>
<td>ðɪs</td>
<td>“this”</td>
</tr>
<tr>
<td>mæθ</td>
<td>“math”</td>
</tr>
</tbody>
</table>

2. Consider the following data from English. According to this data, are [n] and [ŋ] allophones of the same phoneme, or different phonemes?

| mental | “mental” |
| instatʃʌn | “institution” |
| lɪŋɡər | “linger” |
| hændəl | “handle” |
| æŋɡwɔʃ | “anguish” |
| tʃræŋk | “tranq” |
| jən | “yawn” |

3. Consider the following data from Ganda. According to this data, are [l] and [r] allophones of the same phoneme, or different phonemes?

| kola | ‘do’ |
| lwanə | ‘fight’ |
| buulira | ‘tell’ |
| lya | ‘eat’ |
| luula | ‘sit’ |
| omugole | ‘bride’ |
| lumonde | ‘sweet potato’ |
| eddwaliro | ‘hospital’ |
| oluganda | ‘Ganda language’ |
| olulimi | ‘tongue’ |
wulira ‘hear’
beera ‘help’
jjukira ‘remember’
erdyato ‘canoe’
omuliro ‘fire’
effirimbi ‘whistle’
emmeeri ‘ship’
eraddu ‘lightning’
wawaabira ‘accuse’
lagira ‘command’

**Going Beyond**
In English, are [n] and [ŋ] allophones of the same phoneme, or different phonemes? Give data to support your answer. Does your answer support or conflict with your answer to exercise #2? If there is a conflict, how might you explain it?
Determining the Underlying Phoneme

Objective: Given two sounds in complementary distribution and their environments, determine which sound is the underlying phoneme.

Overview
The previous section, “Environments,” described how to decide if two sounds are in contrastive (phonemic) or complementary (allophonic) distribution. If we find that two sounds are different phonemes, then that’s about all we can say about them, but if they are two allophones of the same phoneme, there are still some questions to resolve. One important question is: which is the underlying phoneme? The basic answer to that question is the sound that is in wider distribution. Wider distribution means the sound could occur in more different environments.

To determine distribution, we look at the environment lists—specifically, the side with no overlap. Which list is more variable?

Example Problems

1. This is the right-hand (no-overlap) side of the environments in example #1 (Tongan data) of the previous section (“Environments”). Which sound is in wider distribution?

   | s | t |
   | i | a |
   | i | u |
   | i | o |
   | i | o |
   | i | a |
   | e |

   This is a very clear-cut case. [t] can occur before all sorts of vowels, and [s] can only occur before [i]. Therefore, [t] is in wider distribution and is the underlying phoneme.

2. Compare the following environments. Which sound is in wider distribution, [s] or [z]?

   | s | z |
   | # | ɡ |
   | f | a |
   | t | ɳ |
   | h | i |
   | p | o |
   |  | m |

   In this case, even though [z] looks like it appears with more sounds, all of these sounds can be combined in one category, namely “voiced” sounds (remember that all vowels are voiced). There is no way to group the environment of the [s] sound (because there is no way to group the word
edge with a sound), so [s] is the underlying phoneme, and [z] is the allophone of [s] that occurs in voiced environments.

The key points to remember in identifying the underlying phoneme are:

- **Wider distribution does not mean the list with the most sounds**
- **Wider distribution means the most variable list**
- **Only the side with no overlap has the relevant environment**

**Exercises**

For exercises 2 and 3 of the previous section (“Environments”), state which sound is the underlying phoneme. They are reproduced below for convenience:

1. Consider the following data from English. According to this data, which sound of [n] and [ŋ] is the underlying phoneme?

   mental “mental”
   institution “institution”
   linger “linger”
   handle “handle”

   æŋgwəŋ “anguish”
   tʃræŋk “tranq”
   jən “yawn”

2. Consider the following data from Ganda. According to this data, which sound of [l] and [r] is the underlying phoneme?

   kola ‘do’
   lwana ‘fight’
   buulira ‘tell’
   lyaa ‘eat’
   luula ‘sit’
   omugole ‘bride’
   lumonde ‘sweet potato’
   eddwaliro ‘hospital’
   oluqanda ‘Ganda language’
   olulimi ‘tongue’

   wulira ‘hear’
   beera ‘help’
   jjukira ‘remember’
   cryato ‘canoe’
   omuliro ‘fire’
   effirimbi ‘whistle’
   emmeeri ‘ship’
   eraddu ‘lightning’
   wawaabira ‘accuse’
   lagira ‘command’

**Going Beyond**

1. Explain why we didn’t ask for the underlying phoneme of the first exercise from that section.
2. Consider the Going Beyond question of the previous section (“Environments”): “In English, are [n] and [ŋ] allophones of the same phoneme, or different phonemes? Give data to support your answer. Does your answer support or conflict with your answer to exercise #2? If there is a conflict, how might you explain it?” Does considering the answer to the question provide you with more data which supports or contradicts the decision to label either [n] or [ŋ] as the underlying phoneme?
Rules

Overview
When we find that a pair of sounds constitutes allophones of the same phoneme, the next question is “When do we use which sound?” The general rule is that you start with the basic phoneme, and it turns into a specific allophone version of that phoneme in certain situations. For example, you might say, “/p/ becomes [pʰ] at the beginning of a word,” or more generally, “voiceless stops become aspirated at the beginning of a word.” Formally, these rules would be written as follows:

Symbol Notation:

/p/ → [pʰ] / #_

/p/ becomes [pʰ] when it follows a word boundary (i.e. at the beginning of a word)

Feature Notation:

[-voice, -continuant] → [+aspirated] / #_

Voiceless stops become aspirated when they follows a word boundary
(i.e. at the beginning of a word)

You can see from the examples that the rules take this form:

/phoneme/ → [allophone] / environment,

where the arrow (→) means “becomes” and the slash (/) means “where/when”. So some underlying phoneme becomes some specific allophone when some environment occurs. Your job when writing a rule is to figure out:

• What’s the underlying phoneme and what’s the allophone, and
• What environment triggers the use of the allophone

Feature Notation
Often, it makes more sense to use feature notation than symbol notation for rules. This allows you to see precisely what is changing. For example, if /k/ is showing up as [g] in a certain environment, you know right away that the change is in voicing, since [g] is simply the voiced version of [k].
Example Problems

1. In the “Environments” section, example problem #1, we looked at data from Tongan, and decided that [s] and [t] were in complementary distribution. Here are the environments:

<table>
<thead>
<tr>
<th></th>
<th>s</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>_i</td>
<td>i</td>
<td>_i</td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>_i</td>
<td>e</td>
<td>_i</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>#</td>
<td>a</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>_a</td>
<td>o</td>
<td>o</td>
<td>_o</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

Write a rule in symbol notation which describes when each allophone occurs.

Remember that in “Determining the Underlying Phoneme,” example problem #1, we decided that [t] was the underlying phoneme since it occurs in more variable environments (before [a], [u], [o], or [e], while [s] only occurs before [i]). Since [t] is the underlying phoneme, the rule will look like this:

/\textipa{t}/ \rightarrow [s] / \text{some environment}\

What environment do we want to put there? Well, we want to say when [s] occurs. Looking at the data, it’s pretty clear that [s] only occurs before [i]. So that’s what we write.

/\textipa{t}/ \rightarrow [s] / _i\

How would we write the rule for this example in feature notation? We’ll have to consult our old phonetics charts to figure out the relevant features of [t], [s], and [i].

/\textipa{t}/ : \{+alveolar, -voice, -continuant\} \\
/\textipa{s}/ : \{+alveolar, -voice, -continuant\} \\
/\textipa{i}/ : \{+high, -back, +tense\} \\

Note that when we write the rule, we only have to include those features of allophone [s] that are different from the features of the underlying phoneme /\textipa{t}/.

\{+alveolar, -voice, -continuant\} \rightarrow \{-continuant\} / _\{+high, -back, +tense\}\

2. The English plural “-s” can be pronounced as either [s] or [z] depending on the context. Here’s some data:

<table>
<thead>
<tr>
<th>hobbits</th>
<th>habits</th>
<th>trolls</th>
<th>tfrolz</th>
</tr>
</thead>
<tbody>
<tr>
<td>elves</td>
<td>elvz</td>
<td>wraiths</td>
<td>re@s</td>
</tr>
<tr>
<td>orcs</td>
<td>orks</td>
<td>ponies</td>
<td>poniz</td>
</tr>
<tr>
<td>humans</td>
<td>hjumonz</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

34
Write the rule (in feature notation) for when [s] appears and when [z] appears.

First, we write the environments as per the previous section.

\[
\begin{array}{ll}
&s & Z \\
t_# & v_# \\
k_# & n_# \\
\theta_# & l_# \\
& i_#
\end{array}
\]

We’re concerned with the sound on the left (obviously, since the right side of both lists is composed entirely of word boundaries.) Because [z] occurs with both vowels and consonants and [s] occurs only with consonants, [z] appears to be in wider distribution. Therefore we could write the rule in symbol notation as follows:

\[
/z/ \rightarrow [s] / \{t, k, \theta\}_
\]

Now it’s just a matter of converting the rule to feature notation. First of all, what feature, if any, is common to \{t, k, \theta\} but not \{v, n, l, i\}? Consulting our chart, we see that \{t, k, \theta\} are all unvoiced. Interestingly enough, voicing also accounts for the difference between [z] and [s].

\{+alveolar, +voice, +continuant\} \rightarrow \{-voice\} / \{-voice\}_

**Exercises**

Consider the exercises in the previous section (“Determining the Underlying Phoneme”). Can you write the rule for these exercises? Remember to write in both symbol and feature notation.

1. Consider the following data from English. According to this data, what is the rule that predicts the occurrence of the [n] vs. the [ŋ] sound?

\[
\begin{array}{ll}
\text{mental} & \text{æŋwɔʃ} \quad \text{anguish} \\
\text{institution} & \text{tʃræŋk} \quad \text{tranq} \\
\text{linger} & \text{jɔn} \quad \text{yawn} \\
\text{handle} & \text{hændal}
\end{array}
\]
2. Consider the following data from Ganda. According to this data, what is the rule that predicts the occurrence of the [r] vs. the [l] sound?

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>kola</td>
<td>‘do’</td>
</tr>
<tr>
<td>lwana</td>
<td>‘fight’</td>
</tr>
<tr>
<td>buulira</td>
<td>‘tell’</td>
</tr>
<tr>
<td>lya</td>
<td>‘eat’</td>
</tr>
<tr>
<td>luula</td>
<td>‘sit’</td>
</tr>
<tr>
<td>omugole</td>
<td>‘bride’</td>
</tr>
<tr>
<td>lumonde</td>
<td>‘sweet potato’</td>
</tr>
<tr>
<td>eddwaliro</td>
<td>‘hospital’</td>
</tr>
<tr>
<td>oluganda</td>
<td>‘Ganda language’</td>
</tr>
<tr>
<td>olulimi</td>
<td>‘tongue’</td>
</tr>
<tr>
<td>wulira</td>
<td>‘hear’</td>
</tr>
<tr>
<td>beera</td>
<td>‘help’</td>
</tr>
<tr>
<td>jjukira</td>
<td>‘remember’</td>
</tr>
<tr>
<td>eryato</td>
<td>‘canoe’</td>
</tr>
<tr>
<td>omuliro</td>
<td>‘fire’</td>
</tr>
<tr>
<td>effirimbi</td>
<td>‘whistle’</td>
</tr>
<tr>
<td>emmeeri</td>
<td>‘ship’</td>
</tr>
<tr>
<td>eraddu</td>
<td>‘lightning’</td>
</tr>
<tr>
<td>wawaabira</td>
<td>‘accuse’</td>
</tr>
<tr>
<td>lagira</td>
<td>‘command’</td>
</tr>
</tbody>
</table>

**Going Beyond**

1. In example problem 2, when we developed a rule for plural pronunciation, we decided [z] was the basic underlying phoneme. Given what the rule turned out to be, what’s an alternate explanation for why [z] is in wider distribution than [s]? Could [z] and [s] actually be equally ranked (neither one more “underlying” than the other, or you can’t tell)?
Rule-writing Conventions

Symbols To Know
/X/  X is a phoneme (underlying sound)
[Y]  Y is an allophone (actual pronunciation in this environment)
-> “becomes”; “shows up as”
/  “when”; “where”; “in the environment of”
_  placeholder for the allophone, showing its location in the environment
C  consonants
V  vowels
Ø  null/nothing
#  word boundary (can be used for the beginning or end of a word)
$  syllable boundary

Basic Rule Format
/phoneme x/ -> [allophone y] / _ [trigger environment z]
/phoneme x/ becomes [allophone y] when it comes before [trigger environment z]

/phoneme x/ -> [allophone y] / [trigger environment z] _
/phoneme x/ becomes [allophone y] when it comes after [trigger environment z]

/phoneme x/ -> [allophone y] / [trigger environment z] _ [trigger environment α]
/phoneme x/ becomes [allophone y] when it comes between [trigger environment z] and [trigger environment α]
Analyzing Data Sets for Phonemic and Allophonic Distinctions

Objective: Given a data set, identify whether a pair of sounds represent different phonemes or allophones of the same phoneme in that language. If different allophones, write the rule.

Step-by-Step Guide
Each one of these steps is described in detail a previous section.

Question:
Consider the following data from Daga. Are [s] and [t] allophones of the same phoneme or of separate phonemes?

jamosivin ‘I’m licking’
simura ‘whisper’
jamotain ‘they will lick’
asi ‘grunt’
atu ‘little’
anet ‘we should go’
senao ‘shout’
ut ‘hit’
senao ‘we should go’
topan ‘old’
senao ‘shout’
urase ‘hole’
siuran ‘salt’
siuran ‘salt’
tuian ‘I kill’
wagat ‘holiday’

Step 1: Write the environments
Make one list for each sound of interest. List every individual environment (preceding and following sound, separated by a place-holding underscore) in which that sound is found. Use a pound sign (#) to represent word boundaries (the beginning or end of a word.)

<table>
<thead>
<tr>
<th>s</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>o_i</td>
<td>o_a</td>
</tr>
<tr>
<td>a_i</td>
<td>a_#</td>
</tr>
<tr>
<td>#_e</td>
<td>a_#</td>
</tr>
<tr>
<td>a_e</td>
<td>o_u</td>
</tr>
<tr>
<td>#_i</td>
<td>#_o</td>
</tr>
<tr>
<td>u_e</td>
<td>#_a</td>
</tr>
<tr>
<td>#_u</td>
<td></td>
</tr>
</tbody>
</table>

Step 2: Compare the environments for overlap (one side at a time)
- Is there overlap on the left? Yes, because both the list for [s] and the list for [t] have [o], [a], # on the left side of the underscore.
- Is there overlap on the right? No, because the list for [s] has [i] and [e] on the right side of the underscore, while the list for [t] has [a], #, [u], [o]

Step 3: Make the call: phonemes or allophones
- Overlap on both sides means the sounds are in contrastive distribution and are phonemes.
- No overlap on either side means the sounds are in complementary distribution and are allophones of the same phoneme.

In this case, we have one side with no overlap, so the sounds are in complementary distribution and we can conclude they are allophones.
**ALLOPHONES ONLY: Step 4: Write the rule**

We need to fill in a rule of the form  
/underlying phoneme/ → [allophone version] / environment

**Step 4a. Decide which allophone is the underlying phoneme.**
The allophone that is in wider distribution is the underlying phoneme. To determine distribution, we look at the side of the environments where we found no overlap.

<table>
<thead>
<tr>
<th>Sounds following [s]</th>
<th>Sounds following [t]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>[a] [o]</td>
</tr>
<tr>
<td>[e]</td>
<td>[u] #</td>
</tr>
</tbody>
</table>

Right away, we can see that [t] is in wider distribution because while [s] only comes before certain vowels, [t] comes before certain vowels and word boundaries.

**Step 4b. Write the rule.**
We already have all the information we need to write the rule in symbol notation.

**Symbol Notation**

/\t/ → [s] / _ [i, e]  
“/\t/ becomes [s] before [i] or [e]”

For many rules, especially those with multiple sounds, it will make more sense to write at least part of the rule in feature notation. What’s another way we can say “[i] or [e]”? Consult your vowel chart. What do they have in common (that [a], [o], and [u] do not have in common)?

**Partial feature notation**

/\t/ → [s] / _ [-back, -low]  
“/\t/ becomes [s] before non-back, non-low vowels”

We can also put the whole rule into feature notation by describing [t] and [s] in terms of their features. First we must uniquely describe [t], then we can describe [s] in terms only of the feature(s) that make it distinct from [t].

(Note that because we are working with another language, there may be additional sounds/features to take into account. These will usually be provided for you if they are relevant. For the purposes of this problem, we will assume the features we developed for English are sufficient.)

**Full Feature Notation**

/+alveolar –voice –continuant/ → [+continuant] / _ [-back, -low]  
“voiceless alveolar stops become continuant before non-back, non-low vowels”
Another Example Question
Consider the following data set from Tongan. Are [m] and [t] different phonemes, or allophones of the same phoneme (if they are allophones, provide the rule)?

tauhi  “to take care”
sisi    “garland”
motu    “island”
mosimosi “to drizzle”
motomoto “unripe”
fesi    “to break”
sino    “body”
totonu  “correct”
pasi    “to clap”
fata    “shelf”
movete  “to come apart”
misi    “to dream”

Following the step-by-step guide:
Step 1: Write the environments.

\[
\begin{array}{ll}
\text{t} & \text{e}_e \\
\text{#}_a & \text{m} \\
\text{o}_u & \text{#}_o \\
\text{o}_o & \text{i}_o \\
\text{#}_o & \text{o}_o \\
\text{a}_a & \text{#}_i
\end{array}
\]

Step 2: Compare environments for overlap.
When we look at the left side of each sound, we notice that there is overlap – they share [#, o].
When we compare the right side of each sound, we notice that there is overlap there too – they share [o].

Step 3: Decide on phonemes or allophones.
Since there is overlap on both sides (contrasting distribution), this part is easy – these two sounds are distinct phonemes in this language.

Exercises
1. Consider the following data from English. According to this data, are [n] and [ŋ] phonemes or allophones (if allophones, provide the rule in symbol and feature notations):

mental  “mental”
instatʃən  “institution”
lɛŋɡər  “linger”
hændəl “handle”

æŋɡwəʃ “anguish”
tʃræŋk “tranq”
jən “yawn”
2. Consider the following data from English (use the IPA transcriptions, not the standard orthography glosses.) According to this data, are [θ] and [ð] allophones of the same phoneme, or different phonemes (if allophones, provide the rule in symbol and feature notations)?

boθ “both”
ðoʊz “those”
ɹæðər “rather”
θım “thin”
ʃaðərZ “fathers”
θɪŋk “think”

3. Consider the following data from Ganda. According to this data, are [r] and [l] phonemes or allophones (if allophones, provide the rule in symbol and feature notations):

ekola ‘do’
lwana ‘fight’
buulira ‘tell’
lya ‘eat’
luula ‘sit’
omugole ‘bride’
lumonde ‘sweet potato’
eddwaliro ‘hospital’
oluğanda ‘Ganda language’
olulimi ‘tongue’
wulira ‘hear’

beera ‘help’
jjukira ‘remember’
eryato ‘canoe’
omuliro ‘fire’
effirimbi ‘whistle’
emmeeri ‘ship’
eraddu ‘lightning’
wawaabira ‘accuse’
lagira ‘command’

4. Consider this data from Tongan, a Polynesian language. Are [s] and [t] different phonemes or allophones of the same phoneme (if allophones, provide the rule in symbol and feature notations)?
tauhi “to take care”
sisi “garland”
motu “island”
mosimosi “to drizzle”
motomoto “unripe”
hesi “to break”
sino “body”
totonu “correct”
pasi “to clap”
fata “shelf”
movete “to come apart”
misi “to dream”

vii
**Going Beyond**
You may run into a problem at a certain point in this data set. Consider this data from Tojolabal. Are [t] and [tʰ] separate phonemes or allophones of the same phoneme? If allophones, what rule specifies their distribution?

\[
\begin{align*}
tʃítam & \ 'pig' & \ tʃatʰ & \ 'kind of plant' \\
makton & \ 'a patch' & \ mutʰ & \ 'chicken' \\
potot & \ 'kind of plant' & \ nahatʰ & \ 'long' \\
tinăn & \ 'upside down' & \ tinatʰ & \ 'seed'
\end{align*}
\]

Note: The ejective /tʰ/ is a separate phoneme and therefore will be considered as just another sound, not relevant to solving this problem.\(^{viii}\) (Also, remember that /tʃ/ is one sound!)
Phonological Processes

Objective: Given a rule, determine what, if any, phonological process is at work.

Overview & Examples

Some phonological rules are examples of specific phonological processes.

Assimilation and Dissimilation

Assimilation is when a sound changes to be more like its environment; dissimilation is when a sound changes to be less like its environment.

Recall that rules are of the format:

/underlying sound/ → [surface sound] / environment

Therefore, in an assimilation rule, the surface sound (allophone) and environment will have some feature in common which the underlying sound lacks.

Example of an assimilation rule:

{+alveolar +continuant –voiced} → {+voiced} / {+voiced}_

In a dissimilation rule, the underlying phoneme and the environment will share some feature which the surface sound lacks.

Example of a dissimilation rule:

{+alveolar +continuant –voiced} → {+voiced} / {-voiced}_

Epenthesis and Deletion

So far we have looked at rules where a particular environment triggers speakers to use a particular allophone version of an underlying phoneme. Certain trigger environments can also lead to epenthesis (insertion of an additional phoneme) and deletion (dropping a phoneme in some environment).

Rules With Ø (Deletion & Epenthesis)

[X] → Ø / Y _ Z

“X is deleted between Y and Z”

Ø → X / Y _ Z

“X is inserted between Y and Z”

Examples

1. Consider the following words in English (pronounce them together as you would in normal speech):

   “fast task”
   “camp fire”
   “first grade”
   “second semester”
You may notice that some consonants are dropped (some may not be, depending on dialect). So, a lot of people pronounce these words in the following manner:

<table>
<thead>
<tr>
<th>underlying representation</th>
<th>surface pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“fast task”</td>
<td>[fæst tæsk]</td>
</tr>
<tr>
<td>“camp fire”</td>
<td>[kæmp fæjr]</td>
</tr>
<tr>
<td>“first grade”</td>
<td>[først gred]</td>
</tr>
<tr>
<td>“second semester”</td>
<td>[sekənd semɛstər]</td>
</tr>
</tbody>
</table>

What happens here is that the medial consonant in a cluster ([t] or [d] in this case) gets deleted in surface pronunciation. This is the process of deletion. The rule for this would be:

$/t, d/ \rightarrow \emptyset / C \_ C$

2. Consider the following words in Spanish:

- escuela  “school”
- esperar  “to wait”
- esfera   “sphere”
- estructura “structure”

And the following loanwords from English:

- estandar “standard”
- esmoquin “smoking jacket”

What happens to the English words when they are borrowed into Spanish? It looks as if they are made to comply with the Spanish standard rules of pronunciation (which require an initial “e” [ɛ] before a sC cluster). So, an “e” is inserted to comply with Spanish pronunciation. This is the process of epenthesis. The rule for this would be:

$\emptyset \rightarrow [ɛ]/ \#_sC$

3. Consider the following words from English (pay attention to the vowels):

- jab    jam
- sag    sang
- dad    Dan

You may notice that the vowel [æ] in the first column of words is slightly different from the same vowel [æ] in the second column. The [æ] in the second column is nasalized. The reason for the nasalization becomes clear when we look at the environments. The environment on the left side is the same, so the difference must be on the right. In the first column, the sounds on the right are all voiced oral stops; in the second column, the sounds are all nasal stops. So, the vowel assimilates to its environment (becomes nasalized when it’s next to a nasal sound):

$/æ/ \rightarrow [+ nasal] / _ [+nasal] C$
4. Consider the following data from Kitharaka\textsuperscript{ix} (\textgamma{} is a voiced velar fricative):

\[
\begin{array}{ll}
/\text{ka} + \text{pandi}/ & \rightarrow [\text{yapandi}] \quad \text{a small grasshopper} \\
/\text{ke} + \text{tu\textgamma{}}/ & \rightarrow [\text{yetu\textgamma{}}] \quad \text{a pole} \\
/\text{ka} + \text{cuma}/ & \rightarrow [\text{ycuma}] \quad \text{a small piece of iron} \\
/\text{ke} + \text{kundi}/ & \rightarrow [\text{ye\textgamma{kundi}}] \quad \text{a group} \\
\end{array}
\]

What happens when a [ke] or [ka] prefix is added to a word that starts with a voiceless sound? The [k] in the prefix becomes voiced, so that it is different from its environment. Notice here that the vowel between the two consonants does not seem to play a role. This is sometimes the case with phonological data sets – sometimes, the environment that triggers a change is not the immediate environment. So, in this case of dissimilation, the rule is:

\[
/k/ \rightarrow [+ \text{voice}] / _V [-\text{voice}]C
\]

Here, it is the voicing that is important, so we ignore, in partial feature notation, the change from \textregular -continuant to +continuant that also occurs.

\textbf{Exercises}

1. Consider the following data from a dialect of English sometimes spoken in Eastern Massachusetts (“:” next to a vowel indicates a long vowel); what phonological process could account for the transformation from underlying to surface form? Provide the rule:

<table>
<thead>
<tr>
<th>underlying form</th>
<th>surface form</th>
<th>orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>[sə:ɪŋ]</td>
<td>[sə:ɪŋ]</td>
<td>sawing</td>
</tr>
<tr>
<td>[aɪdɪə]</td>
<td>[aɪdɪə]</td>
<td>idea</td>
</tr>
<tr>
<td>[wɑʃəntən]</td>
<td>[wɑʃəntən]</td>
<td>Washington</td>
</tr>
<tr>
<td>[kləθ]</td>
<td>[kləθ]</td>
<td>cloth</td>
</tr>
</tbody>
</table>

2. Consider the following data from English; what differences in pronunciation do you find among the basic word and the word with an attached suffix? Pay specific attention to the last sound in the root word. What phonological process could account for the transformation? Provide the rule:

<table>
<thead>
<tr>
<th>Root</th>
<th>Root + morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>hand</td>
<td>handed</td>
</tr>
<tr>
<td>long</td>
<td>longer</td>
</tr>
<tr>
<td>limb</td>
<td>limber</td>
</tr>
</tbody>
</table>
3. Consider the following data from English; keep in mind that all of these words have the same underlying prefix: “con-”. What phonological process could account for the transformation from underlying to surface form of the prefix “con-”? Provide the rule:

- conjoin commerce
- contraction compose
- contemplate comply
- conspire combine
- configure
- conclude

4. Consider the following data from English; keep in mind that all of these words have the same underlying suffix: “-al”. What phonological process could account for the transformation from underlying to surface form of the suffix “-al”? Provide the rule:

<table>
<thead>
<tr>
<th>Noun</th>
<th>Adjective</th>
<th>Noun</th>
<th>Adjective</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>personal</td>
<td>pole</td>
<td>polar</td>
</tr>
<tr>
<td>region</td>
<td>regional</td>
<td>circle</td>
<td>circular</td>
</tr>
<tr>
<td>autumn</td>
<td>autumnal</td>
<td>single</td>
<td>singular</td>
</tr>
<tr>
<td>cause</td>
<td>causal</td>
<td>uvula</td>
<td>uvular</td>
</tr>
</tbody>
</table>