What is Phonology?

John A. Goldsmith
Contents

Chapter 1. Alternations 5
1. Sounds and spelling 5
2. Flap 1 6
3. Poor workmanship: s and z 7
4. th 7
5. Flap 2 8
6. Summarizing 10
7. Recognizing the same thing 11
8. Summarizing again 14
9. Flap 3 15
10. The case of today...tonight...tomorrow 18
11. On words 19

Chapter 2. Pure phonology 23
1. The tradition of pure phonology 23
2. Distribution 25
3. Phones with distinctly limited distribution: part 1 26
4. Some more examples 30
5. Free Variation 34
6. Summarizing 36
7. Limited Distribution, Part 2 36
8. Section to be moved to Chapter Five (Word structure)? 40

Chapter 3. Syllables 45
1. Sonority 54
2. Rethinking the syllable: strict syllables 56

Chapter 4. Patterns in language 59
1. Introduction 59
2. Local 59
3. Zero-order generalizations 61
4. First-order generalizations 62
5. Second order generalizations 63
6. Distinctive features, markedness, and underspecification 64
7. Patterns in diphone sequences 66
8. Important theoretical digression 75
CHAPTER 1

Alternations

...it is requisite that each word contain in it so many distinct characters as there are variations in the sound it stands for. Thus the single letter a is proper to mark one simple uniform sound; and the word adultery is accommodated to represent the sound annexed to it in the formation whereof there being eight different collisions or modifications of the air by the organs of speech, each of which produces a difference of sound, it was fit the word representing it should consists of as many distinct characters thereby to mark each particular difference or part of the whole sound.¹


1. Sounds and spelling

PHONOLOGY is the study of the sound systems in language; studies, being what they are, aim to provide us with methods of analysis—which in this case means the analysis of spoken utterances which provide us the means to represent them on paper in a way that provides us with a deeper insight into how each language works. The reader who comes to this book with no knowledge of phonology has a double handicap: in the first place, the handicap of knowing nothing of phonology (a problem that we hope to do something about quite soon), but at the same time the potential handicap of already knowing rather well an old and not very systematic method of analyzing the sounds of English and representing them on paper: standard, written English, which we call English orthography (I will use that term as a synonym for spelling).

It would be pointless for me to ask of you to turn off your knowledge of English orthography as you enter into the arena of phonology, for we can no sooner turn off that knowledge than we could turn off our ability to maintain our balance as we walk down the sidewalk. All we can do is take our knowledge of written English and try to step back from it; we can try to open our ears and really listen to what it is that we say and what it is that we hear all around us.

¹ The works of George Berkeley, vol 1. (ed Luce and Jessop). First published 1709; revised 1732.
I will assume that you are a speaker of English, and that you can produce various sounds out loud, and that you will do your best to hear them as you say them—or in some cases, that you imagine as best you can how other speakers of English pronounce words. As you do that, you will find that you need to make different and often finer distinctions than the standard spelling system of English permits. That awareness will be a sign of increasing phonological sophistication.

I daresay we all have some recollections, dim as they may be, of being taught to read. The teacher taught us the connections between sounds and letters, and soon we came to see the connections between sequences of letters and whole words. Now we must go back and think about the sounds themselves, and not presume that our letters do a complete and accurate job of representing what we say and what we hear. It is not at all unlikely that the reader has internalized various beliefs about linguistic sounds and how they relate to spelling that will have to be jettisoned.

2. Flap 1

Let's begin with a rather tricky case, one which will take us quite some time to unravel. I will suppose that you speak a standard and familiar dialect of American English. You notice one day that in your pronunciation of a word such as *lettuce*, the sound that you utter when you expect to produce a *t* is quite different from the *t* that you produce in *tea* or *telephone*. Do say these words out loud, and attend to how you pronounce them. Something is odd about the sound in *lettuce*, it sounds a bit like a *d*, and as you think about it some more and make a few observations, you notice that when you say the word *potato* out loud, the two *t*'s sound quite different. The first is a real *t*—whatever that might mean; at least, it's much the same as the *t* in *tea*. But the second *t* in *potato* is that odd little sound, the same as the one you make in *lettuce*. Why do you speak this way, you might just ask yourself?

Phonologists, the people who study phonologies, have given a name to the sound which we shall explore: they refer to it as a *flap*, and it differs from the sound that begins the word in *tea*, which is a *stop*, for it truly stops the flow of air through the mouth for a brief period. There is a symbol used to represent the sound of the flap: it is a ⟨r⟩, and to emphasize that when we speak of the flap we are referring to the sound, we often put square brackets around it: [ɾ]. The stop "t" is symbolized by the letter "t", but in order to avoid unfortunate ambiguities, when we mean to refer to the sound, we will explicitly put the "t" between

\footnotesize{We can do better than that, actually. If you have access to a computer linked to the Internet and the World Wide Web, you can listen to the actual sounds of the pronunciations I discuss in this book. Find this information at this URL: xxxx[}
brackets: [t]. So: the letter that we write as \( t \) is sometimes pronounced as a stop [\( t \)] and sometimes as a flap [\( [t] \)]. Why is that?

3. Poor workmanship: s and z

It would likely occur to us rather quickly that we should at least consider the possibility that English has a poor spelling system, and inexplicably uses the same letter ("t") to represent two different sounds. This sort of thing can happen. We can find many cases where English orthography turn out to be confusing, certainly, and perhaps confused. The letter s often represents both the sound of the \( s \) in sink, but also the similar z-sound in zinc, as it does frequently when surrounded in the spelling by vowels (as in wise) or at the end of a word (as in lies). We refer to the difference between the \( s \)-sound and the \( z \)-sound as one of voicing: \( z \) is voiced, while \( s \) is not. English is consistent in its spelling at least to the point where all written \( z \)’s in true English words are pronounced as [\( z \)] (that is, as voiced sounds), but \( s \)’s are an often unpredictable bunch: the written \( s \) can represent either the sound \( s \) or the sound \( z \), depending on quite a few things.

If English had no letter \( z \), and we used \( s \) (let’s ignore \( c \) in this, which can be pronounced [\( s \)], as in nice) both in words with the \( s \)-sound and the \( z \)-sound, then we would have a case where the language just ignored in its spelling system a distinction that is found in the sounds. Could the pronunciation of our \( t \) sometimes as a [\( t \)] and sometimes as a [\( [t] \)] be something like that?

It’s tempting to think so, though eventually we will see that this is not the case. We’ll see, in fact, that the question about the relationship between the sounds [\( t \)] and [\( [t] \)] has nothing to do with spelling at all. But it’s important to pursue the question of orthography for a while, because as we get started in this business of phonological analysis, spelling and pronunciation is pretty much all that we have to hold on to. (It’s also true that people who have never thought about linguistics have some very odd ideas about the relationship between linguistic knowledge and our spelling system, ideas that stand up very poorly against any rational thought, and some of what we will do early on is to confront and disarm some of these ideas.)

4. th

There is another case that we might look at where one written form stands for two quite different sounds. The one I have in mind involves not a single letter, however, but the pair of letters (the term often used for that is “digraph”) th. The pair of letters th rarely is used to represent the sequence of sounds represented individually by \( t \) and by \( h \) (though sometimes it does, as in the word cathouse); it usually is used to represent one of two sounds. One is the sound found in thin, thing, math, catharsis, and many other words; the other is the sound found in
words such as thy, the, those, writhe, either, bother, and many others. These two sounds are quite different, and phonologists have different names for the two. The sound in thin or math is said to be voiceless, just as the sound [s] is; the symbol [θ] is used for this sound. The sound in thy or either is voiced, and the symbol [ð] is used for that sound.

[θ] and [ð] are quite different sounds, even though they are both represented in our writing system by “th”. That these two sounds are quite different is supported by the observation that it is not hard to find words that we know are different that differ only by having [θ] in one and [ð] in the other. Thy and thigh differ (despite what the spelling seems to suggest) solely in the voicing of the first sound: thy begins with [ð], while thigh begins with [θ], and similarly, either and ether differ not in their vowels (as the orthography, again deceiving us, seems to suggest) but in their middle consonant. Either sports a [ð], but ether has a [θ]. How do we know which sound to use—[θ] or [ð]—in any given word? There are some rules of thumb that might be helpful, like verbs that end in -the all end in the sound [ð]. But when all is said and done, the spelling system that we have in English simply makes no serious effort to represent the difference between these two sounds, [θ] and [ð]. And that might give us some reason to take seriously, once again, the possibility that the sounds [t] and [ð] are likewise two distinct sounds not represented by our standard orthography.

5. Flap 2

But even if that were so (and it’s not), it would not put an end to the question that began our account: why is there a flap in lettuce? The sounds of lettuce are no different from the sounds of the phrase let us ..., as in let us begin! If we wanted to mark the sound there, we would have to write le[t] us begin. We have another case of a [t] that’s being represented orthographically by a t. Fine; but a sticky question arises here, for while we say the word lettuce with a flap every time (at least we Americans do—of course the British don’t, but they don’t ever have a flap represented by a t), the same can’t be said of the word let. If we say that word alone (phonologists say, “say the word in isolation”), we find two pronunciations used by speakers of English, but neither of them contain a flap [t]; they both end on a different sound.

We need now to be a bit more careful in our description of our t’s. I have drawn your attention to the difference between [t] and the sound at the beginning of the word tea, which I described as a [t], and this difference is rather clear once you begin to pay attention. But the initial sound in tea, in English, is relatively unusual in a certain respect; it is pronounced with a puff of air, called aspiration, following the release of the tongue (which is to say, after the tongue pulls away from the top of the mouth). Almost anyone who is a native speaker of English who has tried to learn another language (such as French or
Spanish) has been faced with the challenge of producing a t without this aspiration. When we wish to make it clear that a t is being pronounced with aspiration, we indicate that with a following raised h, like this: [tʰ].

The word let, when said in isolation, can end with either what linguists call a “non-released” t, written [t] (and this is the more common American pronunciation), or a released t, written [tʰ]. What are these sounds, and how do they differ? A released [tʰ] (which is often used by speakers from New York, for example, for a t that comes at the end of a sentence) has a burst of air that is released after the complete closure of air flow that is created by the tongue to make the sound [tʰ]. After that closure has been held for a brief period, the tip of the tongue comes down a bit from the top of the mouth, and a brief burst of air flows over the top of the tongue.

The alternative pronunciation of the t is as a glottalized sound. Here too the tip of the tongue comes up to the roof of the mouth, but the flow of air up from the lungs is closed at the vocal cords — in the throat — and so even if the tip of the tongue comes down from the roof of the mouth, the outward flow of air has been checked down in the throat, and so there is a complete silence following the closure made by the t. This is the sound we have called [t]: it is unreleased, because the glottis closes, and by the time it opens again, the tongue is no longer making the gesture of a t.

We will need a symbol for these t-sounds. The more common glottalized t is represented thusly: [tʰ], while the released t is represented by the symbol [t]. For now, we will focus on the more common pronunciation of t as [t]. We can use these terms to summarize our observations regarding the pronunciation of the word let. When the t of let comes at the end of a sentence (or more generally, a phrase), it is pronounced as a glottalized t [tʰ]. When let is followed by the word us, the t is pronounced as a flap [ɾ].

Is the t of let pronounced as a flap regardless of what word follows let? The answer is No. Let’s consider a range of words that might follow let, and observe how the t of let is pronounced in these cases; see Table 1.

There appears to be a principle lying behind the decision to use the unreleased t and the flap (though you may object to my calling that a “decision”). No matter how long we extend this list, we will find that the principle at work is this: when let is followed by a word starting with a vowel, its t is pronounced as a flap [ɾ], and when let is followed by a consonant (that is, anything else), the t of let is pronounced as a [t]. This is a correct generalization, but a few additional points must be borne in mind.

First of all, why is there a flap in let him in the house, if him starts with a consonant (“h”)? It’s not hard to see that the h of him is not
1. Alternations

| 1. Unreleased [t'] |  
|-------------------|---|
| let go            |   |
| let Mary go       |   |
| let Paul go       |   |
| let Tom go        |   |

| 2. Flap [r]       |  
|-------------------|---|
| let a man go free |   |
| let a boy go home |   |
| let him in the house | |
| let Amy do it    |   |

Table 1. Ts

pronounced in speech at a normal conversational style – and this is true regardless of whether ts or flaps are at issue. In a phrase like help him do it, the h of him is not present at normal speech rates. Only if we slow down considerably, pronouncing each word as a separate mini-phrase does the h of him reappear. So when we say that the t of let becomes a flap before a vowel, we really mean in front of a vowel, regardless of whether there is a consonant in the orthography or not.

Second, it is always possible to put some emphasis on the word let (which in the cases we are looking at is the main verb of the sentence), and the result of that stress is that for the purposes at hand, let is treated as a separate phrase, with a bit of the lengthening which is the telltale sign of the word in question being treated as if it were at the end of a phrase. We will come back and talk about this at length.

6. Summarizing

What we have done so far is this. We have chosen a particular word (let) that happens to end in a t, and we have varied the right-hand environment in which that word finds itself. We have noticed that the final sound of let is realized in one of two ways (either glottalized or flapped), and that the decision as to which form is used is not arbitrary, but rather is based on a simple phonological principle: glottalized before a consonant, flapped before a vowel.

Our next step is going to be to see whether the word let is in any way special or unusual in this regard, or whether any word that ends in a t will display this same kind of behavior, realizing the t as unreleased or flapped depending on whether a vowel or consonant follows. But before doing that, I would like to explain why there is something that might once have been controversial about our exploration so far. In what we have done so far, we have leaned very heavily upon our knowledge of English, in the sense that we all know the simple word let, and we can easily build sentences using the word, as in (1) and (2). But
in order to have arrived at a point where we could easily carry out these procedures, we had to have many years of instruction in the English language. For most of us, that instruction was informal, on our parents’ knees and in the streets, but instruction it was nonetheless. We internalized the rules of English to the point where without considerable care and attention, we cannot attend to the sounds: we effortlessly pass on from the sounds to the words. So we have internalized the principles that determine whether the \( t \) is realized as glottalized or appeded—we have internalized that rule, and our task right now is to make explicit as best we can what that rule is. But the sticking point is this: it is our tacit knowledge of that rule which makes it possible for us to identify the word let as being one and the same word in all the different contexts in which it appears.

### 7. Recognizing the same thing

This point may be difficult to appreciate at first, but it was extremely important in the development of what is called structuralist phonology, or phonemics. It is still extremely important, but it is less frequently appreciated, and we will consider this point of view very seriously in the next chapter. But we can make the following observation now. When we undertake various phonological observations, we will always be trying to in some fashion or other look at the same “thing” from several points of view—such as when we looked at the final sound of the word let in different contexts, depending on what word follows. How do we know that it is the same “thing” that we are looking at in the various cases? In the one case we have considered so far, there would seem to be nothing controversial: the word let is the word let, and we know it when we see it. Or so it would seem; but much more difficult cases lie in wait for us. I will just mention one example now, and we will deal with many others later, in due time.

Suppose we compare the words five and fifth. Are they related? Our immediate reaction is probably, Yes, they’re related, just as six and sixth are related, and seven and seventh. But what about one and first, or two and second, or even three and third? The phonological form of the words of each pair (one/first, two/second) is very different, and no one, I daresay, would say that the words of each pair are related in a sound-like way. There is no correspondence of any sense between the sounds of one and first: the \( n \) of one does not correspond to any particular sound of first, and if one is actually pronounced much like the word won is, that has no effect on how first is pronounced. The same could be said for the pair two/second, and if we know something about the history of English, we’ll even be aware that two is a word that goes far back in the history of the language, while second is a word that was brought in from French. In these cases, the words in each pair are related, undoubtedly, but not related in a phonological way (that is,
as far as the sounds are concerned). What about *three/third*? Here we see some phonological relationship: they both start with [θ], and both have an *r* in them — but other than that, nothing more connects them, and just as importantly, there is no pattern of relationship between the sounds which appears in any other pair of words that we would be inclined (on grounds of meaning) to relate to each other.

What, now, about the case I mentioned first, *five/fifth*? If we do decide that the two forms are related — in a phonological, sound-based way — does it follow that there is a single “thing” that they share in common, and which we can identify? There certainly is a great deal that these two words have in common. If we take *fifth* and lop off the -*th* suffix, we get *fif-* , and we will immediately note three things about it. First, it differs from *five* in that it ends in an *f*, not a *v*, but *f* and *v* are pretty similar — they are identical except that *f* is voiceless and *v* is voiced; second, although we use the same letter to mark the vowel in *five* and in *fif-* , the two vowels are quite different in pronunciation. The vowel of *five* is what we were taught in school was a long vowel, while the vowel of *fif-* is a short vowel. Still, there seems to be a regular relationship between these two different vowel sounds, a point we will look at in detail later on. And third, the *fif-* that we get when we lop off the suffix -*th* is also found in two other apparently related words, *fifteen* and *fifty*.

The case of *five/fifth* is the first case that illustrates for us that it may be difficult to know whether in fact we are keeping something the same when we compare two pronunciations. Don’t misunderstand: to do any kind of analysis, we must compare various versions of “the same thing”; but it can be extremely difficult, in the worst cases, to know if it really is “the same thing” that we’re comparing at a given moment.

To draw an analogy: suppose we are analyzing photos taken by a spy plane (or satellite) of enemy territory, and we need to know what changes have taken place since the last time we took photos. We may have thousands of photos from each pass; but these photos are of use to us only if we know how to match them up with the photos taken on a previous pass, so that we can note the differences. What allows us to be sure that two photos are photos of the same place? Sometimes it will be easy to be sure: there may be a famous statue in the photo, or a unique waterfall, but in many other cases it will be hard to know just what we’re looking at. That’s a bit like the problem we face here, and it is one that we will come back to time and again.

To return, then, to the question we posed a moment ago: is there anything special about the word *let*, or would any word ending with a *t* show the same variety of ways of realizing that *t*? In (3) I have given some phrases that include a word ending in *t*, and followed by either a vowel or a consonant. We see that in each case, a word-final *t* is pronounced as a flap [ɾ] when the word is followed by a word that
begins with a vowel, and it is realized as a glottalized t [tʰ] otherwise. “Otherwise“ means here that either a word follows that begins with a consonant, or no words at all follows.

(3) examples

We are in a position now to try to make a phonological generalization. It is phonological because it will refer only to phonological entities, though for now we will have to be a bit vague as to just what that means. But for present purposes, we may focus on the fact that the generalization does not need to refer to any particular word, like *let*, or *home*, or *him*, etc – just sounds, like [tʰ] and [ɾ].

(4) Generalization: When an English word ends in a t, that t is realized as a flap when a word immediately follows which begins with a vowel; otherwise the t is realized as a glottalized stop.

It might be helpful to formulate this generalization in a more graphical fashion. We might start thusly:

```
  t
 /\  /
[ɾ] [ɾ]
```

[ɾ] [tʰ] before a word in all other cases beginning with a vowel

We should ask ourselves at this point what the diagram in (4) is actually displaying. The [ɾ] and the [tʰ] we can simply say are descriptions of sounds. We will call these kinds of entities phones. And the “t” that sits above them, connected by a line to each of them? What is that t?

We’re not in a position to answer that question firmly at the moment. Three possibilities come to mind immediately:

- it could really be a [ɾ] (which happens to be converted into, and realized as, a [tʰ] in some cases);
- really a [tʰ] (which happens to be converted into, and realized as, a [ɾ] in some cases);
- or something else that is different from both a [tʰ] and a [ɾ].

And that something else – it might be either yet some third phone, one which we have not described yet; or – and this turns out to be the right answer – something more “abstract“, something which is not pronounceable per se, but which can be realized phonetically in various ways.

This notion lies at the heart of phonology, and various names have been given to it, each name carrying with it a great deal of intellectual baggage; of these, the most common are phoneme and underlying
1. Alternations

We will be dealing with these phonemes and underlying segments throughout the course of this book.

8. Summarizing again

Let us review what we have done so. We have suggested that there is a single phoneme (\(t\)) which can be found at the end of various words (let, hit, pet...), among other places no doubt, and which will be realized as either a \([t']\) or a \([c]\) – not freely, but in a fashion governed by the generalized given in (4). Do bear in mind that the entire force of the argument derives from the notion that we would like to say that each particular word of English has a single phonological specification – a specification as a list of phonemes, roughly speaking. If we didn’t care about that, we could say that there are twice as many words (from a phonological point of view) than we used to think that there were. The word spelled let, for example, would have two phonological specifications: [let] and [let'], each used in a particular phonological context (one before vowels, the other elsewhere).

These two approaches are not as different as they may sound. The first one (which really is a better way of putting it, as we will eventually see) says that there is a single underlying form for let, made up of three phonemes, the last of which is a \(t\) which can be realized in one of two ways. The second approach says that there are two ways that let can be realized: either as [let'] or as [ler]. But the second approach must add a further statement: all words which are realized with a final [\(c\)] before a vowel will be realized with a final [\(t'\)] elsewhere. This generalization is an entirely empirical generalization: we can search hundreds, even thousands of words, and this generalization will rarely be violated if at all, with the most solid kind of statistical results that anyone could ask for. In a formulaic shorthand, we can say that the existence of \(X[\(c\)]\) before vowels implies the existence of \(X[\(t'\)]\) elsewhere, where \(X\) stands for any string of phones. Or better yet, we can restate it this way: a word in English may take the form

\[
X\{\{c\}, \{t'\}\},
\]

where \(X\) is any string of segments, and the final segment is a tightly linked pairing of the sounds [\(c\)] and [\(t'\)], realized according to the principle we have stated several times now (and furthermore, no words can end in [\(c\)] without also having a kindred form ending in [\(t'\)], and vice versa). But in this final statement of the generalization, there seems to be no real difference between saying that there is an abstract phoneme \(t\) which can be realized as a [\(c\)] or a [\(t'\)], on the one hand, and saying that English words come in pairs summarized in a formula like \(X\{\{c\}, \{t'\}\}\).
She took a test
(both ts followed by a stressed vowel) – flap impossible.
She ate ten cookies.
She brought the ten largest cookies.
The tedium can be overwhelming.

UNSTRESSED: flap [ɾ] impossible here as well.

True telepathy might be a boon to humanity.
He met with a tenacious opponent in the final round.
I tried to buy a tomato.
I saw Toledo from the air (or, I saw Topeka...)
It’s not always easy to construct a topology for a fifteen-dimensional sphere.

Figure 1. Example 6

We might say, as many phonologists working in the structuralist tradition did say, the phoneme \( t \) is just a way of referring to the pairing \( r, t^\prime \) along with its rule of realization.

Given the apparent equivalence of these two ways of talking about things (though the equivalence will eventually collapse, as we look at more complex cases), we will adopt the first way of speaking about those phonemes or underlying segments.

9. Flap 3

example 6:

Our generalization in (2) tells us how a \( t \) is pronounced when it comes at the end of an English word. What about the other cases – what if the \( t \) come at the beginning of a word, for example? Let’s construct a list of a good number of words that begin with \( t \), followed by various stressed and unstressed vowels.

(Mind: I’m leaving four words out on purpose for now – the words to, tonight, today, and tomorrow; we’ll come back to them later. Yes, they would have flaps in them in many of these sentences, unlike all of the other \( t \)-initial words we are looking at here.)

In none of the phrases in (6) will we find a flap [ɾ], and that hold true regardless of the stress, or lack of it, on the vowels before and after the \( t \). So next to generalization (4) is (3), rather different:

(2) [repeated] Generalization: When an English word ends in a \( t \), that \( t \) is realized as a flap when a word immediately follows which begins with a vowel; otherwise the \( t \) is realized as a glottalized stop.
1. Alternations

<table>
<thead>
<tr>
<th>Previous vowel:</th>
<th>Following vowel:</th>
<th>Stressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstressed:</td>
<td>Sanity, audacity</td>
<td>Italian</td>
</tr>
<tr>
<td>Stressed:</td>
<td>Italy, writing</td>
<td>Detail, Retail, Beethoven, atoll</td>
</tr>
</tbody>
</table>

(3) Generalization: When an English word starts with a t, that t is realized as a [tʰ], not as a flap [ɾ].

These two generalizations focus on ts that are word-final and word-initial; what of t’s that are neither, but are rather word-internal (or as we say, word-medial)? There are easily hundreds, even thousands of words to look at; t is one of the most common phoneme in the language, after all [check its relative frequency to that of i]. If we listed everyone, we would find many of them with a flap, and many with a [tʰ], and we would have to spend a good deal of time sorting out the two groups.

Let’s focus first on those words where the t is surrounded on both sides by vowels, and let’s then divide that group into four, based on the stress of the vowels on either side. Since we may speak of vowels as being either stressed or unstressed, that gives us four groups:

*(But: Latin, button, satin, Martin)*

Why do I divide the data up in this way? Only because I have been looking at this data for years: there is no prior reason that I can give, other than this classification works. And it works in the sense that the flapping properties of the t’s in each of these categories is consistent: the t’s in (8a) may optionally be flaps or optionally be true [ɾ]’s; the t’s in (8b) and (8d) cannot be flaps, while those in (8c) must be flaps.

Putting that together, then, we can say (4), with one special case to which we will return immediately, that of button:

(4) When a t is word-internal and surrounded by vowels, it must be realized as a flap [ɾ] when the preceding vowel is stressed and the following vowel is unstressed; it may be realized as a flap [ɾ] when the vowels on either side are unstressed; otherwise, it must be realized as a true [t].

We could put it slightly differently, though making precisely the same point:

(5) When a t is word-internal and surrounded by vowels, it can be realized as a flap only if the following vowel is unstressed. In that case, it may be a flap if the preceding vowel is unstressed, and must be a flap otherwise (i.e., if it is stressed).

However, when the t is immediately followed by an unstressed vowe which in turn is followed by an n, as in button, then an additional complexity arises. Some speakers, including many from the South, have a flap in this environment; others, including this writer, have a more complex articulation here. In producing the t, the tongue comes
up to the roof of the mouth, making the gesture of a true \( t \), but the glottis closes (as we earlier observed it would do when the \( t \) comes at the end of a word). Then something unusual happens, or rather two somethings simultaneously. The glottis opens, allowing air to flow up from the lungs, and at the same time the velum — the gateway from the back of the mouth to the nose — opens up, allowing the air to rush out of the nose rather than through the mouth, the mouth still being blocked by the tongue placed at the roof of the mouth. This new sound is thus an \( n \), which is what we have when air flows through the nose rather than the mouth, and the blade of the tongue is closes off the flow of air through the mouth.

This peculiar realization of the \( t \), occurring when an unstressed vowel followed by \( n \) is what follows the \( t \), has only a small domain in which it is operative, though many of the words that are involved in it are extremely common (like button), and the realization of the \( t \) in this glottalized fashion has precedence over flapping for many speakers, including this writer.

Using the generalizations in (9/9'), we can cast our net a bit wider, and ask what principle governs the realization of all the other word-internal \( t \)s. Making sure to avoid compound nouns (which function differently) like ant eater, we find that no additional flaps come to light: all the flaps that we find occur when the following vowel is unstressed, but the nature of the consonants neighboring the \( t \) makes a difference.

\begin{enumerate}
\item If any consonant immediately follows the \( t \), then we cannot have a flap \([t]\). If the following consonant is an \( r \), the \( t \) and \( r \) together make a sound not all that different from the sound of \( ch \); the sound is certainly not that of a flap, but it’s not a true \([t]\) either: words like trick, troop, Petri, paltry.
\item If an \( r \) precedes the \( t \), the flap is normal, with one special case. The normal cases include words like artichoke, Sparta, Jakarta, article, artificial, aorta, mortal, and furtive. But just as we noticed earlier, if an unstressed vowel plus an \( n \) follows, the \( t \) will, in the speech of many speakers, be realized as a glottalized \( t \), with a release directly into the \( n \). This is what occurs in such important words as important, though many American speakers (such as President Jimmy Carter, from Georgia, or this writer’s mother, from Minnesota) have a flap even in words like important). present us with clearcut flaps. When a consonant other than \( r \) follows the \( t \), the \( t \) will normally be glottalized, as in Atkins, delightful, platform, beatnik, catnip, atmosphere, etc. When a consonant other than \( r \) precedes, such as an \( l \) or an \( n \), as in altitude or cantaloupe, we generally get a \([t][n]\), though in casual speech, it is true that the combination of \( tt \), and even
more of nt, is produced so quickly that it is not possible to distinguish it from a rapid flap.

No other cases present us with clearcut flaps. When a consonant other than r follows the t, the t will normally be glottalized, as in Atkins, delightful, platform, beatnik, catnip, atmosphere, etc. When a consonant other than r precedes, such as an l or an n, as in altitude or cantaloupe, we generally get a [tʰ], though in casual speech, it is true that the combination of It, and even more of nt, is produced so quickly that it is not possible to distinguish it from a rapid flap.

10. The case of today...tonight...tomorrow

We need to explore one more twist to the story of the flap before we can say that we are done. When we think about words that begin with a t followed by an unstressed vowel, four of the most common words that spring to mind are to, today, tonight, and tomorrow—surely more quickly than tomato or, certainly, Topeka. And yet these four to-words (to, today, tonight, tomorrow) do not follow the generalization that I suggested above for words that beginning with a t: the four to-words do take a form with an initial flap [t]:

(7) We’re going to fly [t]o Seattle on Monday.
What are you going to see [t]onight?
Who will you see [t]omorrow?

The facts regarding the non-flapping of other words starting with t seem to be robust, though. How should we think about the special behavior of the to-words, then?

It won’t do simply to say that their behavior is different because the words are common; while there is a seed of truth to it, we have no particular expectation of what the ways are in which a more common word ought to behave differently from a less common word, so the force of the word “because” is thoroughly mysterious in such a case as this.

We have divided the flap phenomenon up into two pieces, figuring out first how t behaves when it comes at the very end or very beginning of a word: in that case, it either flaps or doesn’t flap (according to the case), with no dependence on stress. Then we looked at how the t behaves when it is inside a word, and there we found that flapping is dependent on an unstressed vowel immediately following the t. In the case of the to-words, the vowel that follows the t is always unstressed, and the t always has the option of being realized as a flap.

It appears, then, that the most economical way to interpret the facts is to conclude that the to-words may maintain a closer phonological relationship with the word to their left; as phonologists say, the to-becomes an enclitic to the preceding word when that word ends with a vowel. The situation is complex, and remains somewhat obscure. For example, in the case of the expression have got to (do something), as in
I’ve got to leave soon, we find a parallel cliticization of the word to the word preceding it. This is indicated in the colloquial spelling I’ve gotta leave soon. What’s striking about that pronunciation is that when we put got and to together, we’d expect two t’s, but in fact there’s only one (from the point of view of sounds) and it turns into a flap: I’ve go[t]a leave soon. This cliticization is all the more striking in view of its absence in otherwise parallel cases. If we say, I forgot to leave, the result is a sequence of two t’s and no flapping; we conclude that there is no cliticization in such a case (and hence to can’t be said to always cliticize to what precedes it): I forgo[t] [t]o leave. With a bit of effort, one can hearly the two t’s, one at the end of forgot and the other at the beginning of to, much like in a compound noun like hot-tub.

This is not quite the end of the story about the flap in English, but we have now seen most of the account. ts that normally are realized as flaps occur at the end of words, or, inside a word, before an unstressed vowel – the flap being obligatory if what precedes is a stressed vowel (followed optionally by an r). And to this, an extra statement about the to-words must be added, as we have just seen.

11. On words

The analysis that we have worked through for the flap in English has employed the notion of “English word” in two different ways, and we would do well to think about these uses. On the one hand, we used our ability to recognize words like let as being one and the same word despite the fact that sometimes it was realized with a final flap and sometimes with a final glottalized t. On the other hand, we have also used the notion of word when we formulated our generalization about the behavior of the flap, for a t will behave differently (regarding its flapping) depending on whether it appears at the beginning, at the end, or in the middle of a word. In this latter case, we need to know where words begin and end (where the word boundaries are, phonologists say) in order for these notions to be applicable in any given case.

Neither of these notions of word is as simple as they might seem at first blush. In fact, these questions sound simple only if you have never worried about them. Let’s slow down for a bit, and consider that first notion, of being able to identify when it is that we have two instances of the same word in a given utterance. Consider the following not so obvious cases.

(8)  a. In the sentence I broke the nail of my left thumb when I hammered in the last nail, are the two words nail the same word? They are both nouns, but their meanings are entirely unrelated.

b. In the sentence, If the paper won’t stick to the wall, stick it in the drawer. are the two words stick the same word?
They are both verbs, but their meanings are only distantly related, if that.

c. In the sentence, *You need to use a nail to nail it down*, are the two words *nail the same word? They belong to different parts of speech: one is a noun, the other a verb, though the meaning of the noun is somehow included in the meaning of the verb.*

d. In the sentence, *I eat what I want to eat*, are the two words *eat the same word? They are semantically closely related, they are both verbs, but one is the third person singular form of the verb, while the other is the infinitive. If we compare this sentence to one based on another verb, we get *I am what I want to be* – are *am* and *be* the same word here? Do we want to treat this last example differently from *I eat what I want to eat*, just because the verb to *be* is (inflectionally) less regular than to *eat*?

e. In the sentence, *I scream, you scream, and he screams for ice cream*, are there any pairs of words (especially the words *scream and screams*) that are the same word? The verbs are all inflected, though for different persons (first, second, and third person singular); and English typically treats first and second person verbs the same, with no suffix added. Does that make the verbs of *I scream* and *you scream* the same, or different? Linguists, employing their vast technical terminology, might well say, The two screams bear different morphosyntactic features, but they are morphologically identical: but what does that mean for the phonologist? Let’s look a bit closer.

Let’s consider the standard approach to this problem – it’s standard because it’s the best we have. It amounts to saying, let’s build a dictionary, pretty much like the dictionary on your shelf, in which we will have separate each entries for each part of speech a word can be used as, and in which only reasonably closely related senses can be combined into a single entry. Each entry in this dictionary will be called a *lexeme* (though nowadays computational linguists use the term *lemma* for this same notion). Each lexeme includes a number of different ‘words’, in the everyday sense – words with different spellings. The entry for woman also contains women; these are two words, but just one lexeme. Write and writing are likewise two words, but they represent just one lexeme, a verb. We’ll say that two words are the same word only if their phonemic representation (which is, of course, somewhat more specific than the spelling) are the same, they represent the same lexeme, and in all of the grammatical specifics that might separate two words, they agree. *I scream and you scream:* five words,
all different; 1st person *scream* and 2nd person *scream* are not the same word, though they represent the same lexeme. (This means that *scream* the noun and *scream* the verb are not even the same lexeme.)

Now we can summarize a little bit better what we have tried to do so far. We have said that we want to find a way to give a single phonological representation to words, plus a set of rules that enable us to understand how and why that single phonological representation is realized differently in different contexts. In view of what we have said about the organization of the lexicon — distinguishing, for example, between lexemes, on the one hand, and the words that fall into various lexemes — we can easily imagine three levels of challenges for phonology. The lowest challenge is to do as we have just said: to find a single phonological representation for each word, and leave it at that. Each word will have its own underlying phonological representation, but (as we have seen for the word *let*) the word may be pronounced differently depending on what words surround it in a given sentence, and phonology will have to come to grips with that fact. Second, we could raise the ante, and ask for an explanation of the relationship between the phonological representations of words contained with the same lexeme. *Write* and *writing*, for example, are two members of the same lexeme; they are separate words, but their phonologies are closely related. *Write* has a [tʰ], while *writing* has a [ʤ]. Perhaps phonology is responsible for that as well. Third, we could raise the ante still higher, and ask for an explanation of the relationship between various lexemes — like the relationship between *hesitate* and *hesitation*, or *five* and *fifteen*.

We will adopt the first of these three as our goal for now. Let us give a name to this task: we will call it the problem of alternation, a word that suggests our interest in the alternative pronunciations that an individual word is capable of. In fact, so that we can talk about the three levels of the problem that we have just defined, we will use the term first order alternations to refer to the different ways individual words will be realized in different contexts (and I will use the word alternations without a modifier to mean this for now); second order alternations to refer to the ways in which the sounds of related words pertaining to a single lexeme are related to each other; and third order alternations to refer to the ways in which the sounds of related words are related to each other.

Phonologists today are broadly in agreement that the matter of first-order alternations is a realistic and attainable goal, and that it falls to phonologists to deal with this problem. We shall later see that phonologists are by and large in agreement (though with notable exceptions) that the second goal is also within the province of phonology as well, though irregularities (which abound in this second area) may stand out as problems that phonology can say little about. The third
area, which deals with the relationship between separate lexemes, especially the phonological relationships, is embattled territory, and we will need a strong background and technology to approach this problem, as we will do later on. Let me emphasize the word alternations. What we shall take to lie at the heart of this concept is the notion that we can grab hold of particular words, identify them for what they are, and look at how their pronunciation varies in different contexts. The term itself is not a new one; it goes back to the structuralist, or phonemicist, framework to which we have till now only alluded, but which will be the focus of the next chapter. Our use of the term alternation is slightly different from the phonemicists' use, though it agrees very much with the spirit of their use of the term. The difference is this: the phonemicists intended to prepare a very constrained and restricted way to determine how phonemes would be realized in utterances, and they did so without making any reference to the problem of alternations as we have defined it. That was their explicit goal. But they recognized that their methods were in a sense quite weak, and would not, in a vast range of cases, provide solutions for the problems of alternations in virtually any language you chose to look at. So they developed the notion of alternation to refer those aspects of the problem of alternations (as we have defined it) which went beyond the methods of phonemes and phonemic analysis.

What this terminological choice means is that we shall be able to ask a question like, Does a certain set of techniques for phonemic analysis provide a satisfactory solution for the problem of alternations in a language – for the different ways in which individual words are realized? Phonemicists would have balked at that formulation of the question, because their use of the term alternation specifically means that part of the problem which could not be handled by phonemic analysis. But our use of the term will be much handier for us, and it is not much out of the spirit of the phonemicists' use. In any event, the term has largely fallen into desuetude, and so we are offering it a new lease on life. Let us turn to the phonemicists' view of the analysis of sounds.