COURSE PACKET

This course packet is intended as a supplement to class lectures and other materials. It contains lots of definitions, background, and general information about how to approach the topics we’re covering, but it certainly doesn’t cover everything we’ll talk about in class – so make sure to pay attention to other handouts, notes you take in class, etc. when doing homework and studying for exams.

TABLE OF CONTENTS

SYLLABUS..............................................................................................................2
  Where to find me..................................................................................................2
  Grading..................................................................................................................3
  Homework............................................................................................................3
  Exams...................................................................................................................4
  Academic honesty.................................................................................................5
  Course website.....................................................................................................6
  Text.......................................................................................................................6
  Attendance............................................................................................................6

1. MORPHOLOGY.................................................................................................7
  1.1. How to do morphological analysis (or any other kind of linguistic analysis)......7
  1.2. Types of morphemes.....................................................................................8
  1.3. Morphological rules.....................................................................................10
  1.4. English prefixes and suffixes.......................................................................11
  1.5. Drawing morphology trees..........................................................................13
  1.6. Parts of speech.............................................................................................14

2. SYNTAX...........................................................................................................16
  2.1. Words, phrases, and constituency tests.......................................................16
  2.2. Phrase structure rules..................................................................................20
  2.3. How to do syntactic analysis (and find phrase structure rules)......................21
  2.4. Drawing syntax trees...................................................................................25
  2.5. Transformations: Syntax rules.....................................................................27
  2.6. Figuring out d-structures..........................................................................31

3. PHONOLOGY....................................................................................................33
  3.1. The sounds of English, and their features....................................................33
  3.2. Minimal pairs...............................................................................................37
  3.3. Allophones and complementary distribution..............................................38
  3.4. Finding the basic phoneme, and phonetic motivation for allophonic changes...41
  3.5. Phonological rules.......................................................................................42
  3.6. Positions in syllables....................................................................................43
SYLLABUS

WHERE TO FIND ME

Class  Monday, Wednesday, Friday 10:10 – 11:00

E-mail  flack@linguist.umass.edu  (e-mail is the best way to get in touch with me)

Course website  http://people.umass.edu/flack/201

Office  South College 313
South College is the old red brick building next to the library.
To get to my office, go in any of the doors on the south side of the building
(the same side as the front door of the library) and go up two flights of stairs.

Office hours  Wednesday 11:00 – 12:30, and by appointment.
Please feel free to drop by during office hours, or to set up an appointment if
you’d like to meet at another time.

Mailbox  In South College 226 (linguistics department office), to the left of the door.
To get to the department office, go in the front door of South College (the
door next to the bike racks, on the east side of the building, facing the library)
and up the stairs; the office is immediately on your left.

Important: Leave homework, etc. in my mailbox, not under my office door.

Important dates for this class

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday, February 19</td>
<td>NO CLASS</td>
<td>Presidents’ Day</td>
</tr>
<tr>
<td>Wednesday, February 28</td>
<td>Midterm 1</td>
<td>Morphology</td>
</tr>
<tr>
<td>Wednesday, April 11</td>
<td>Midterm 2</td>
<td>Syntax</td>
</tr>
<tr>
<td>Monday, April 16</td>
<td>NO CLASS</td>
<td>Patriots’ Day</td>
</tr>
<tr>
<td>(date to be announced later)</td>
<td>Final exam</td>
<td>Phonology</td>
</tr>
</tbody>
</table>
**Grading**

50% homework:

10 assignments (see below)
Each will count for 5% of your final grade.

50% exams:

3 exams (see below)
Each will count for 16.7% of your final grade (that is, 1/3 of 50%, or 1/6 of the total grade).

Extra credit:

You may do up to 2 linguistics experiments for extra credit.

You can sign up to do experiments at [http://wwwx.oit.umass.edu/~linguist/esdb/](http://wwwx.oit.umass.edu/~linguist/esdb/).
(if you don’t want to type all that, find the linguistics department webpage, click on ‘Research’, and click on ‘Experiment Sign-Up’ at the right edge of the page)

Each experiment that you do will replace one of your lowest homework grades.

When you do an experiment, you will receive a slip with your name on it from the experimenter. Bring this slip to me; this is how I know to give you extra credit.

Experiments generally last an hour or less; if you do an experiment that lasts 90 minutes or more, you’ll receive additional credit. Let me know if your experiment was extra-long.

**Homework**

There will be 10 homework assignments; each is worth 5% of your total grade.

Homework is worth half of your total grade in this class, so it’s to your advantage to do it. Exam problems will be very similar to homework problems, and so by doing the homework you’ll be preparing yourself for the exams.

Homeworks will generally (but not always) be distributed on Friday and due the following Friday.

Homework must be in my hands by the end of the class period when it’s due. **No late assignments will be accepted.** E-mailed assignments that show up by class time, or assignments that show up in my box by class time, are fine.

Homework can always be turned in early. You can turn assignments in in class, of course; you can also put them in my mailbox (in South College 226), or send them by e-mail.
If you e-mail work to me you must make sure I receive it. Whenever I receive an e-mail with an assignment from you, I will write right back to let you know I’ve gotten it. If time goes by and you don’t hear from me, there is a good chance that the e-mail gods have eaten your offering before it could reach me. If you think this may have happened, send it again. Keep a copy of any work you e-mail to me, so you can send it again if necessary. If I don’t ever get your e-mailed assignment, this is ultimately your problem, so be careful!

If you think you might not make it to class – you’re not good at alarm clocks, you’re feeling a little queasy, your car broke down, you have a huge paper to write for another class, whatever – you can be safe by e-mailing your homework to me before class time, or even the night before. If you turn up in class it will be a delightful surprise; if you don’t, we won’t have to do the “will you take my homework late” dance.

If your printer only prints in pale pink and yellow stripes until you get a new cartridge – don’t give me pink homework. Really. E-mail it, or handwrite it, or go to the computer lab. A good test – if you’re going to write a little note with a smileyface (🙂) apologizing for the appearance of your homework, don’t hand it to me.

You’re welcome to talk to other students in the class when you’re working on the homework; however, you must write it up on your own, in your own words, using your own examples. Failure to do so will be considered cheating; please don’t cheat (see below on academic honesty).

Exams

There will be three exams:

- Wednesday, February 28: Midterm 1, Morphology
- Wednesday, April 11: Midterm 2, Syntax
- (date to be announced later): Final exam, Phonology

Each exam is worth 16.7% of your total grade – that is, 1/6 of your total class grade, and 1/3 of the 50% that the exams are worth overall.

Everyone in the class will take the exams on the dates announced unless something terrible happens. If you will need to take the exam at some other time, you need to let me know before the scheduled exam time.

The in-class portion of each exam will be designed to be do-able during the regular 50-minute class period; the exams may also include take-home sections. If you need extra time to take an exam, or any other sort of accomodation, let me know ahead of time and that can be arranged.
You may bring a note sheet to each in-class exam. This should be a single piece of 8.5” x 11” paper, with writing on one or both sides. It can be handwritten or typed, and can have anything in the world that you want on it. The only thing you can’t bring is photocopied material from class or printouts from the website. If you make it yourself (you also can’t make one and give copies to your friends, or copy your friend’s), it can be anything. **You may use anything you like on take-home exams.**

Please don’t cheat on exams; see below on academic honesty.

---

**ACADEMIC HONESTY**

**Please don’t cheat.**

Be aware of university regulations on academic honesty; they can be found at [http://www.umass.edu/dean_students/rights/acad_honest.htm](http://www.umass.edu/dean_students/rights/acad_honest.htm) and will be enforced in this class. I assume that you are familiar with this policy.

The following things are considered to be cheating:

- Copying all or part of a homework assignment or exam from someone else.

- Copying homework or exam answers from books, the internet, or anywhere else without acknowledging the source, or when the assignment asks for your own work.

- Cribbing or copying during an exam.

- Helping someone else cheat, i.e. providing your work to someone else for copying.

**If you cheat on a homework assignment or exam, you will get a zero for that assignment/exam. If you cheat on a subsequent assignment or exam, you will fail the course, and will be reported to the Academic Honesty Board.**

Remember, you may discuss homework problems with classmates. You are encouraged to study together for exams. But working together and copying someone else’s work are not the same thing. No part of your homework or exam may be identical to another student’s. Please don’t cheat.
COURSE WEBSITE

http://people.umass.edu/flack/201

What’s there:
- Class syllabus
- Course note packet
- Class handouts
- Homework assignments
- Homework answer keys
- Write-ups of some of the things we do in class
- Exam study guides
- Homework answer keys

Notice that most of these things will also be handed out in class; if you miss class, or lose things, go first to the website and see if you can get the materials you need there before asking me for them.

COURSE TEXT, AND OTHER REFERENCE MATERIALS

There is no textbook for this class. I’ll hand out packets of course notes that you can and should use for reference – they contain lots of basic info about the class material, including lots of definitions, descriptions of what various kinds of rules should look like, etc. Be warned, though, that reading the book is not a substitute for coming to class regularly; it’s intended only as a supplement to lecture. Not every single thing we do in class is covered in the book, and you are responsible for every single thing we do in class (see below on attendance).

ATTENDANCE

Attendance does not directly count towards your grade in this class. Of course, it’s always to your advantage to come to class, if you want to do well – especially because the book doesn’t cover everything we do in class.

Finding out what happened in a class that you missed is your responsibility. If you miss class, check the website for things that may have been handed out; it’s also a good idea to check with a classmate and get an overview of what happened.

If you need to miss a class when homework is due, get the homework to me by class time some other way.

Contact me immediately if you are on an athletic team or other university activity which will require you to be absent from 5 or more meetings of this class. According to university regulations, this will mean that you are not able to take this course.

If you have unfortunate life circumstances that mean you will need to miss lots of class, let me know as soon as possible and we’ll make arrangements accordingly.
I. Morphology

1.1. How to do morphological analysis (or any other kind of linguistic analysis)

Morphology is the study of word formation – how words are built up from smaller pieces. When we do morphological analysis, then, we’re asking questions like, what pieces does this word have? What does each of them mean? How are they combined? Can we predict how they might combine to make other new words?

In general, when you’re asked to do any sort of linguistic analysis, you’ll be given a set of data – words or sentences from some language that generally isn’t, but occasionally is, English – and asked to find patterns in it. Here’s a set of data from Cree, an Algonquian language spoken in Canada – from this, you might be asked something like, how do you say “my (whatever)”?

<table>
<thead>
<tr>
<th>Cree</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>tʃi:ma:n</td>
<td>canoe</td>
</tr>
<tr>
<td>nitʃi:ma:n</td>
<td>my canoe</td>
</tr>
<tr>
<td>so:niya</td>
<td>money</td>
</tr>
<tr>
<td>niso:niya</td>
<td>my money</td>
</tr>
<tr>
<td>wiya:j</td>
<td>meat</td>
</tr>
<tr>
<td>niwiya:j</td>
<td>my meat</td>
</tr>
</tbody>
</table>

Note: Don’t worry about the weird symbols in these words. They’re characters from the International Phonetic Alphabet (IPA) and stand for various sounds; we’ll talk about them in the last section of the course, but for now it doesn’t matter what they mean – just treat them like any other letter.

To figure this out, you should look for pairs of words with small, relevant differences. For example, if you’re trying to figure out how to say that something is “my (thing)”, well, you can look for two words that differ only that one is “my (thing)” and the other is just “(thing)”. The first pair of words are like this – the first is a word for “canoe”; the second is a word for “my canoe”. It seems likely, then, that the difference between the two Cree words corresponds to the difference between the two English phrases. That is, the difference between these two words:

<table>
<thead>
<tr>
<th>Cree</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>tʃi:ma:n</td>
<td>canoe</td>
</tr>
<tr>
<td>nitʃi:ma:n</td>
<td>my canoe</td>
</tr>
</tbody>
</table>

…is just the ni- at the beginning of the second word. So you might think, very sensibly, that adding ni- to the beginning of a word is how you make it “my (thing)”. You can test this hypothesis on other pairs of words whose English translations differ only in the desired way, like these:

<table>
<thead>
<tr>
<th>Cree</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>so:niya</td>
<td>money</td>
</tr>
<tr>
<td>niso:niya</td>
<td>my money</td>
</tr>
</tbody>
</table>
Again, the only difference between the two Cree words is the *ni*- at the beginning, so again, it looks like adding *ni-* to something makes it “my (thing)”. This is confirmed by looking at the last pair of words, which show the same pattern:

\[
\begin{align*}
\text{wiya}: & \quad \text{meat} \\
\text{niwiya}: & \quad \text{my meat}
\end{align*}
\]

So overall, how to do morphological analysis:

When you don’t know anything about the language:

- Find a pair of words whose English translations differ only in a single way that’s relevant to the task at hand.
- Find the corresponding difference in the non-English words – often some letters will be added, or the word will be changed in some systematic other way. Making this change in the non-English word therefore produces the relevant change in the English meaning.
- Check your theory: find another pair of foreign words whose English translations also differ only in this relevant way, and make sure this pair of foreign words change in the same way as the last.

We’ll encounter kinds of morphology that are more complex than this, but this basic method of looking for forms with minimal differences and figuring out how to describe that difference is always a good approach.

1.2. Types of morphemes

**Morpheme** A morpheme is the smallest unit of meaning we have – that is, the smallest piece of a word that contributes meaning to a word.

**Example** The word *trainings* has 3 morphemes in it – *train-ing-s*.

Note: When using a word as a linguistic example, like *trainings*, above, write it in italics (or, if you’re hand-writing, underline it). This makes it much easier to tell which words are your examples and which are your explanations. It’s an important distinction; be careful to make it in all your homework and exams.

To break a word into morphemes, try starting at the beginning of the word and seeing how far into the word you need to go to find a sub-part of the word that has some meaning. For example, in the word *unloved*, the first two letters *un-* are independently meaningful in a way that just the first letter, *u-*, is not – *un-* means something like ‘not (whatever)’, and changes the meaning of the word it attaches to in a predictable way; sub-parts of *un-*, like *u-* or –*n-*, don’t have this property. This means that *un-* is a morpheme.

Once you’ve found the first morpheme, ask yourself whether there’s another meaningful sub-part of the word after that first morpheme. Again, –*love-* is independently meaningful; so is the last
part of the word, -d. So *unloved* has three morphemes: *un*-love-*d*. Some words just have one morpheme, of course – you can’t break down the word *love* into any meaningful sub-parts, for example.

We define different kinds of morphemes based on various properties like where they show up in words.

All morphemes are either free or bound.

**Free**    A free morpheme is one that can stand on its own – that is, it’s an entire word.

*Examples*  *the*, *cat*, *run*, *pretty*, *trapezoid*

Free morphemes may appear with other bound morphemes attached to them; crucially, though, they don’t need to have other morphemes on them.

**Bound**  A bound morpheme cannot stand on its own, but rather must be attached to a free morpheme whenever you say it.

*Examples*  *re*-, *un*-, *-est*, *-er*, *-fer* (see below)

Some morphemes are roots; others are affixes.

**Root**  The primary piece of meaning in a word, to which affixes can be added. In English, a root is often a word itself.

*Examples*  *cat*, *pretty*, *-fer*

A note on bound roots: Most roots are free morphemes – you can say them on their own as words. Some, though, like *-fer* (as in *re*-fer, *in*-fer, etc.), are bound – they have to have other bound affixes attached to them in order to surface. How do we know that *-fer* is a bound root and not just another affix? Well, linguists assume that every word has a root in it somewhere, which gives it its part of speech and things like that. When you look at the parts of *refer*, *re-* is a pretty common prefix – it can apply to all kinds of new words (*re*-napster) and change their meaning in a predictable way – and *-fer* shows up after various other prefixes (*in*-fer, *pre*-fer, *con*-fer, etc.). *-fer* therefore looks like the root of the word *re*-fer. Bound roots are relatively rare in English, but are often more common in other languages.

**Affix**  A morpheme which attaches to roots (or stems), changing their meaning in regular ways.

*Examples*  *re*-, *un*-, *-est*, *-er*, *ing*, *-s*
Stem A form to which an affix is attached. This can be a root, or a root that already has one or more affixes on it but that can still take more.

Examples eat, swim (can take e.g. -er) eater, swimmer (can take e.g. –s)

Affixes are generally either prefixes or suffixes.

Prefix An affix that goes before a root.

Examples re-, un- (re-read, un-loved)

Suffix An affix that goes after a root.

Examples -est, -er, -s (quick-est, quick-er, read-s, book-s)

1.3. Morphological rules

When you’re doing morphological analysis, you’ll be asked to report your results in various ways. Sometimes you’ll be asked to tell whether various morphemes are free or bound, roots or affixes, prefixes or suffixes, etc. Other times, you’ll be asked to write rules that explain how words are built out of morphemes.

The point of writing a rule is to describe exactly what’s going on morphologically in such a way that someone could use your rule to build new words. A good test for whether your rule is right is to try to use it and make sure it gives you the right result. This will become clearer soon.

Morphology rules are sentences that tell you these three (or four) things:

(1) What kind of morphological category you’re expressing.
(2) What change takes place in the stem to express this category.
(3) Where in the stem this change takes place.
(4) Special conditions, if any, on this change (e.g. it might only occur in certain circumstances, on certain kinds of words, etc.).

We can look at a simple example rule that makes the English plural form of a noun (i.e. that takes the root cat and adds an –s to the end to make cats). The rule looks like this:

To make the plural form of a noun, add –s to the end of the noun.

We can break down this rule to show super-explicitly which parts of it are doing which of the four necessary things, like this:
(1) To make the plural form of a noun, add –s to the end of the noun.

Note that in this case (well, for the purposes of our discussion, anyway), the plural is always formed by adding –s, so we don’t need any special conditions – that is, no part (4).

When I talked about testing your rule, here’s what I meant: you’ll write rules based on data – here, the data is the pair of words cat and cats. You can then take the rule and the data and make sure the rule produces the data – so here, you can take the singular form cat and ‘do’ the rule to it – that is, ‘add –s to the end of the noun.’ This produces cats, as it should. This is such a simple, familiar example that testing it seems dumb, but in more complex examples, testing is a great way to make sure you’ve done everything right.

Sometimes it will be harder to write a very simple description of the morpheme being added to the root, because the morpheme might be making a pretty complicated change; even when the thing that happens to the root is complicated, just make sure to explain exactly what happens and your rule will work.

1.4. English prefixes and suffixes

DESCRIBE.

<table>
<thead>
<tr>
<th>AFFIX</th>
<th>MEANING</th>
<th>EXAMPLES</th>
<th>ATTACHES TO A…</th>
<th>FORMING A…</th>
</tr>
</thead>
<tbody>
<tr>
<td>ex-</td>
<td>former…</td>
<td>ex-president, ex-con</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>dis-</td>
<td>not…</td>
<td>dishonest, disloyal, dissatisfied</td>
<td>Adj</td>
<td>Adj</td>
</tr>
<tr>
<td>fore-</td>
<td>…before</td>
<td>foresee, foreshorten, foreshadow</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>in-</td>
<td>not…</td>
<td>incompetent, incomplete, intolerable</td>
<td>Adj</td>
<td>Adj</td>
</tr>
<tr>
<td>mid-</td>
<td>in the middle of…</td>
<td>midseason, midweek, midair</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>mis-</td>
<td>…in a wrong manner</td>
<td>mistake, misunderstand, misspell</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>re-</td>
<td>…again</td>
<td>rework, rethink, reevaluate, redo</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>un-</td>
<td>not…</td>
<td>unhappy, untrue, unsure, unconscious</td>
<td>Adj</td>
<td>Adj</td>
</tr>
<tr>
<td>un-</td>
<td>do the opposite of…</td>
<td>untie, unwrap, uncover, undo, unfold</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>-able</td>
<td>able to be …ed</td>
<td>lovable, fixable, breakable, washable</td>
<td>V</td>
<td>Adj</td>
</tr>
<tr>
<td>-age</td>
<td>the result of …ing</td>
<td>breakage, bondage, dosage</td>
<td>V</td>
<td>N</td>
</tr>
<tr>
<td>Suffix</td>
<td>Definition</td>
<td>Examples</td>
<td>Part of Speech</td>
<td>Position</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>----------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>-(i)al</td>
<td>pertaining to…</td>
<td>national, musical, presidential</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ate</td>
<td>make…</td>
<td>activate</td>
<td>Adj</td>
<td>V</td>
</tr>
<tr>
<td>-ation</td>
<td>act of …ing</td>
<td>relaxation, meditation, realization</td>
<td>V</td>
<td>N</td>
</tr>
<tr>
<td>-dom</td>
<td>state of being…</td>
<td>wisdom, freedom, boredom</td>
<td>Adj</td>
<td>N</td>
</tr>
<tr>
<td>-en</td>
<td>make…</td>
<td>gladden, widen, soften, roughen, redden</td>
<td>Adj</td>
<td>V</td>
</tr>
<tr>
<td>-er</td>
<td>one who …s</td>
<td>baker, teacher, owner, wanderer</td>
<td>V</td>
<td>N</td>
</tr>
<tr>
<td>-ful</td>
<td>full of…</td>
<td>graceful, joyful, playful, hopeful</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-hood</td>
<td>state of being a…</td>
<td>sisterhood, childhood, neighborhood</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>-ic</td>
<td>pertaining to…</td>
<td>organic, atmospheric</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ify</td>
<td>make (into a)…</td>
<td>classify, objectify, solidify</td>
<td>N/Adj</td>
<td>V</td>
</tr>
<tr>
<td>-ion</td>
<td>act or result of …ing</td>
<td>protection, compensation, reflection</td>
<td>V</td>
<td>N</td>
</tr>
<tr>
<td>-ish</td>
<td>like a…</td>
<td>boyish, childish, foolish, sheepish</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ity</td>
<td>the quality of being…</td>
<td>sanity, activity, passivity, masculinity</td>
<td>Adj</td>
<td>N</td>
</tr>
<tr>
<td>-ive</td>
<td>tending to…</td>
<td>assertive, comprehensive, reflective</td>
<td>V</td>
<td>Adj</td>
</tr>
<tr>
<td>-ize</td>
<td>make…</td>
<td>visualize, unionize, crystallize</td>
<td>N/Adj</td>
<td>V</td>
</tr>
<tr>
<td>-less</td>
<td>without…</td>
<td>penniless, priceless, hopeless</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ly</td>
<td>like a…</td>
<td>friendly, womanly, manly, cowardly</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ly</td>
<td>in a … manner</td>
<td>slowly, happily, hurriedly, foolishly</td>
<td>Adj</td>
<td>Adv</td>
</tr>
<tr>
<td>-ment</td>
<td>act or result of …ing</td>
<td>adjournment, government, movement</td>
<td>V</td>
<td>N</td>
</tr>
<tr>
<td>-ness</td>
<td>quality of being…</td>
<td>happiness, firmness, kindness</td>
<td>Adj</td>
<td>N</td>
</tr>
<tr>
<td>-ous</td>
<td>characterized by…</td>
<td>famous, poisonous, rancorous</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-ship</td>
<td>state of being (a) …</td>
<td>championship, kinship, governorship</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>-some</td>
<td>characterized by…</td>
<td>troublesome, burdensome, worrisome</td>
<td>N</td>
<td>Adj</td>
</tr>
<tr>
<td>-y</td>
<td>…like</td>
<td>mealy, pulpy, mousy, icy, fruity, fiery</td>
<td>N</td>
<td>Adj</td>
</tr>
</tbody>
</table>
1.5. Drawing morphology trees

In addition to rules, we use tree drawings a lot in linguistics to represent the internal structure of words, sentences, etc. Trees are diagrams that look sort of like mobiles, where all the parts of e.g. the word hang off the bottom, and they’re grouped together two at a time until you’ve joined together all the parts of the word.

A tree for a word containing two morphemes, like cat-s, is really simple:

```
N
\   \nN   s
|   |
cat
```

This shows that the noun cats has two pieces: the root cat (which is a noun) and the suffix –s. Cat is underneath an N; this shows that cat is a noun. –s isn’t under a part of speech label because it doesn’t have a part of speech on its own. The top N in the tree – representing the whole word cats – is made by joining the N above cat and the –s; there’s a N where these two join together because the word they make – cats – is still a noun.

Trees are more illuminating when we’re combining more than two morphemes at once. We’ve said in class that if you’re building a word like unloved, with three morphemes, we need to combine them two at a time – first you make loved, and then you add the suffix un- to it. (remember that doing it the other way around would make the non-word unlove first, and that would be a problem). We can show the order that morphemes combine in a tree, like this:

```
Adj
\   \  
un-  Adj
\   |   
V   -ed
|   |
love
```

Again, this tree has the root love under a part of speech label V. Love and the suffix –ed join to form an adjective, denoted by the Adj at the top of the lines connecting V and –ed. The new adjective loved combines with the prefix un- to form yet another adjective, unloved.

Important properties of morphology trees:
- Roots are always directly below a part of speech label.
- All words created by adding an affix to a root (or stem) are labeled with their part of speech.
- Affixes are added to stems one at a time. That is, no part of speech label has more than two lines below it.
### 1.6. Parts of speech

You’ve probably heard definitions for parts of speech like this: “A noun is a person, place, thing, or idea” or “A verb is an action word.” That’s lovely, but they’re slippery definitions – we generally agree that a word like *appetite* is a noun, but it’s not really a person, place, thing, or idea; similarly, *seem* is a verb, but it’s not really an action word. So instead of these meaning-based definitions of parts of speech, in this class we’ll use structural definitions – that is, definitions based on the structure of a word, and/or its position in a sentence structure.

<table>
<thead>
<tr>
<th></th>
<th><strong>NOUN</strong></th>
<th><strong>VERB</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affixes it may have</strong></td>
<td>-er/-or owner, actor</td>
<td>-ate designate, appreciate</td>
</tr>
<tr>
<td></td>
<td>-ity brevity, solemnity</td>
<td>-ify terrify</td>
</tr>
<tr>
<td></td>
<td>-ment government</td>
<td>-ize unionize</td>
</tr>
<tr>
<td></td>
<td>-ness happiness</td>
<td>-en darken, lighten</td>
</tr>
<tr>
<td></td>
<td>-(t)ion vision, rendition</td>
<td>en- enroll, ennoble</td>
</tr>
<tr>
<td><strong>Affixes it can take</strong></td>
<td>plural –s chairs, appetites</td>
<td>past –ed played</td>
</tr>
<tr>
<td></td>
<td>possessive –’s chair’s, appetite’s</td>
<td>present –ing playing</td>
</tr>
<tr>
<td><strong>Syntactic position</strong></td>
<td>May follow a determiner, or a determiner and an adjective: the (big) chair</td>
<td>May follow to, as an infinitive to play, to terrify</td>
</tr>
<tr>
<td></td>
<td>my (first) car</td>
<td>May follow an auxiliary</td>
</tr>
<tr>
<td></td>
<td></td>
<td>should play, can play</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th><strong>ADJECTIVE</strong></th>
<th><strong>ADVERB</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affixes it may have</strong></td>
<td>-able readable, edible</td>
<td>Usually has the suffix –ly.</td>
</tr>
<tr>
<td></td>
<td>-ed frightened</td>
<td>slowly, occasionally, terribly</td>
</tr>
<tr>
<td></td>
<td>-ish childish</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ive defective</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-y sandy, hairy</td>
<td></td>
</tr>
<tr>
<td><strong>Affixes it can take</strong></td>
<td>comparative –er taller, shorter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>superlative –est tallest, shortest</td>
<td></td>
</tr>
<tr>
<td><strong>Syntactic position</strong></td>
<td>May appear between a determiner and a noun: the big chair, my first car</td>
<td>May appear after he did it he did it slowly</td>
</tr>
<tr>
<td></td>
<td>May appear after seemed: the chair seemed big the wolf seemed hairy</td>
<td>Cannot appear after seemed <em>the chair seemed slowly</em></td>
</tr>
</tbody>
</table>
There are a few other English parts of speech that we’ll talk about:

**Determiner**

**Syntactic position**  Before a noun (and adjective, if one precedes the noun); only one determiner can precede a single noun (*the a car*).

*The cat sat on my favorite chair.*

**Examples**  *the, a, my, your, his, her, its, our, their, this, that, those, some, all, every, one, two, three...*

**Auxiliary**

**Syntactic position**  Before a verb; no more than three auxiliaries may appear before a single verb.

*I could have been lying on the beach right now.*

**Examples**  *be (is/am/are/were/being...), have (has/had/having...), can, could, may, might, will, would, shall, should, must*

**Pronoun**

**Syntactic position**  In a position normally occupied by an entire noun phrase

*I want you to go to him and get it tomorrow.*

**Examples**  *I, me, mine, we, us, ours*

*you, yours*

*he, him, his, she, her, hers, it, its, they, them, theirs, one*

**Preposition**

**Syntactic position**  Before a noun phrase; usually only one preposition can precede a single noun phrase (*on above the desk*).

*Before the ice age, dinosaurs wandered across the earth.*

**Examples**  *about, above, across, after, against, among, around, ago, as, at, before, behind, below, beside, between, but, by, despite, down, during, for, from, in, inside, into, off, on, out, over, past, since, than, through, to, toward, under, until, up, with, without*
2. **Syntax**

2.1. **Words, phrases, and constituency tests**

Sentences are made up of words, but we can show that there’s another level of organization in between – that is, that sentences are made up of phrases, and phrases are made of words. Syntax is concerned with the ways that words are assembled into phrases, and phrases are assembled into sentences. This section will describe what kinds of groups of words count as phrases, and how to identify them.

We can see a relationship between words and phrases by noting than whenever you find a single noun in a very simple sentence, you can replace it with a group of words, all related to each other and crucially containing a noun, and the sentence will still be grammatical. So for example, starting with the simple sentence:

John runs.

…we could replace the single noun *John* with a group of related words:

The skinny little guy in the neon tank top runs.

…and the sentence is still fine – and, if you know the John I know, it means the same thing. This group of words is called a noun phrase – it’s a phrase that’s built around a central noun (*guy*), and the whole phrase can be traded for a single noun and the whole sentence will still make sense. Notice that while we can replace this noun phrase (abbreviated ‘NP’) with a noun and still have a good sentence, we can’t replace it with any other part of speech (parts of speech are noted):

*Swim* (V) runs.
*Fast* (Adj) runs.
*Quickly* (Adv) runs.
*Down* (P) runs.

We’ll also talk about two other kinds of phrases: verb phrases (VP) and prepositional phrases (PP). A verb phrase is a group of words that can substitute for a single verb, or that can have a single verb substituted for them; a prepositional phrase is – unsurprisingly – a group of words that can substitute for a single preposition, or that can have a single preposition substitute for them.
In the simple sentence *John runs*, where *runs* is the verb, we can put in various groups of words – crucially each including a verb – in place of *runs*; these phrases are VPs:

> John runs through.
> John runs quickly.
> John runs to the store, over the bridge, through the forest, past the ice cream stand, and then back home.

We can see that prepositions also have a relationship with prepositional phrases: anywhere you can have a preposition on its own, like in a simple sentence like *John runs up*, you can have a longer group of words including a preposition – this group is, of course, a PP:

> John runs up the street.
> John runs through town.

When we talked about figuring out what part of speech a single word was, we said that while intuitive definitions like “A noun is a person, place, or thing” might be useful guides, to be really precise we needed to use structural definitions of categories like nouns to figure out whether a given word was a noun. Similarly, while an intuitive definition of a noun phrase like “A noun phrase is a group of words that describe a noun, and can go in sentences where nouns can go” may be a useful guide, we’ll use formal tests called *constituency tests* to prove whether groups of words are particular kinds of phrases.

**Constituency test** A way of testing whether a group of words forms a phrase. There are different constituency tests to check whether groups of words are different types of phrases.

The constituency test for noun phrases (NPs) is the *pronoun test*, where you replace a group of words that you think might be a noun phrase with a pronoun; if the resulting sentence is grammatical, and means the same thing as the original sentence (that is, the resulting sentence is a shorter way of saying the original one), the group of words is a NP.

**Pronoun test** Replace a group of words with a pronoun, like *he, she, him, her, it, they*, or *them*. If any of these replacements results in a grammatical sentence with the same meaning as the original sentence, the group of words is a noun phrase. If none of the replacements has this result, the words are not a noun phrase.

**Examples** I played the *smallest violin in the world* → I played it

The new sentence is grammatical and means the same thing as the original sentence, so the italicized words in the first sentence form a noun phrase.

Most people love *Ashton Kutcher* → Most people love him

The new sentence is grammatical and means the same thing as the original sentence, so the italicized words in the first sentence form a noun phrase.
I played the smallest violin in the world $\rightarrow$ *I played the it the world

The new sentence is not grammatical, so the italicized words in the first sentence do not form a noun phrase.

Bob threw *away his homework  $\rightarrow$  Bob threw it

The new sentence is grammatical, but it does not mean the same thing as the original sentence, so the italicized words in the first sentence do not form a noun phrase.

This constituency test is basically just a single, standard way of doing the kind of substitutions we were talking about at the beginning of this section. It’s just a little more rigorous, substituting with a very particularly defined set of words, which allows us to all do the same substitutions and prove things in the same way. Consistency is a linguist’s best friend.

Note: Phrases can be just single words. We can prove this using the pronoun test on a sentence like Bob likes cheese $\rightarrow$ Bob likes it. Since cheese can be replaced with the pronoun it, cheese is a one-word NP.

There’s also a constituency test for VPs: the do-form test.

**Do-form test**  Replace a group of words with a form of do so. If the replacement results in a grammatical sentence with the same meaning as the original sentence, the group of words is a verb phrase. If the replacement does not have this result, the words are not a verb phrase.

**Examples**  I’m going to talk to your mother $\rightarrow$ (I said I’d talk to your mother, and) I’m going to do so.

The new sentence is grammatical, so the italicized words in the first sentence form a verb phrase.

Bob talks to his mother every day $\rightarrow$ (Jim talks to his mother once a month, and) Bob does so every day

The new sentence is grammatical, so the italicized words in the first sentence form a verb phrase.

Jim wants to talk to his mother $\rightarrow$ *Jim wants do so mother

The new sentence is not grammatical, so the italicized words in the first sentence do not form a verb phrase.
Finally, the constituency test for prepositional phrases is the there/then test.

**There/then test**  Replace a group of words with *there* or *then*. If the replacement results in a grammatical sentence with the same meaning as the original sentence, the group of words is a prepositional phrase. If the replacement does not have this result, the words are not a prepositional phrase.

**Examples**  I want to go to the moon  →  I want to go there.

The new sentence is grammatical, so the italicized words in the first sentence form a prepositional phrase.

I want to eat a huge sandwich on the day after tomorrow  →  I want to eat a huge sandwich then.

The new sentence is grammatical, so the italicized words in the first sentence form a prepositional phrase.

I want *to* go to the moon  →  *I want there

The new sentence is not grammatical, so the italicized words in the first sentence do not form a prepositional phrase.

I have to admit that the there/then test is kind of crappy. While you can reliably identify every single NP by finding a good-sounding pronoun substitution, and every single VP with a good-sounding do-form substitution, there are some PPs where just no substitution of *there* or *then* will sound particularly good. In these cases, I offer two guidelines which will help you find PPs anyway:

- Every P is part of a PP (just like every N is part of an NP and every V is part of a VP; see below). Whenever you find a P in a sentence, it will be part (or maybe all) of a PP. This reliable fact is the best way to start trying this occasionally frustrating test.

- Once you find the P, and you’re trying to figure out what (if any) of the words around it comprise the PP, start trying to trade *there* or *then* for every possible set of words close to and including the P. Eventually, there will be some substitution that sounds significantly better than the rest. The resulting substituted sentence may not have the same meaning as the original sentence, but at least it will sound grammatical. When a *there/then* substitution sounds relatively good like this, take that as an indication that you’ve found the group of words that form the PP.
One (well, three) huge thing(s) to remember

**EVERY N IS PART OF A NP.**
**EVERY V IS PART OF A VP.**
**EVERY P IS PART OF A PP.**

This means that, when I ask you to find all of the phrases in a sentence, if the sentence has 3 nouns, two prepositions, and a verb (like the one below), you should find 3 NPs, 2 PPs, and a VP.

Try to find all six phrases in this sentence:

Mark fled from the scary old haunted house on the hill.

Embodied in this guideline is the fact that phrases often contain other phrases – for example, both PPs and two of the NPs are inside the VP in the *haunted house* sentence, and in fact each PP has an NP inside it.

### 2.2. Phrase structure rules

In studying the structure of sentences, a major goal is to provide a general description of what sentences, NPs, VPs, PPs, etc. look like in different languages. We do that by writing general rules for the structure of each of these kinds of phrases, and calling these rules *phrase structure rules*.

Let’s say we’ve figured out (as we likely have by now) that every single English sentences has a noun phrase and then a verb phrase. We can write a schematic rule that says this – that thusly describes the structure of all English sentences – like this:

\[ S \rightarrow NP \ VP \]

If you were to read that rule in words, it would sound like “A sentence consists of a noun phrase followed by a verb phrase.” To be really explicit, the pieces of this rule mean the following:

- **S** = “a sentence”
- **→** = “consists of”
- **NP** = “a noun phrase” (this is old news)
- **VP** = “a verb phrase” (also old news)

…and in the part of the rule after the arrow, left-to-right order in the rule corresponds to left-to-right order in a sentence, and also in a syntax tree. That’s where the “noun phrase followed by a verb phrase” part comes from – the order of the constituents after the arrow.
Syntax rules can, of course, build things other than sentences, like this:

\[ \text{NP} \rightarrow \text{Adj } \text{N} \]

This rule says that a noun phrases consists of an adjective followed by a noun, like in *pretty girls*. We know, though, that not all English noun phrases have adjectives in them – you can start a sentence with just a noun, as in *Girls laugh*, *Mary laughs*, etc. Other times, you can put a whole bunch of adjectives in a noun phrase: *Big ugly girls laugh*, etc. There are symbols that go in rules to denote optionality (i.e. when something isn’t necessary) and the possiblity of multiple instances of something:

\[ \text{NP} \rightarrow (\text{Adj*}) \text{ N} \]

This rule says ‘A noun phrase consists of one or many optional adjectives followed by a noun.’ Parentheses around something mean that it’s optional; an asterisk after it mean that you can have a bunch in a row.

\[
\begin{align*}
\text{NP} & \quad = \quad \text{“a noun phrase”} \\
\rightarrow & \quad = \quad \text{“consists of”} \\
(\text{Adj*}) & \quad = \quad \text{“one or more optional adjectives” (this can be broken down further:)} \\
(\text{Adj}) & \quad = \quad \text{the adjective is optional} \\
\text{Adj*} & \quad = \quad \text{one or more adjectives} \\
\text{N} & \quad = \quad \text{“a noun”}
\end{align*}
\]

One way to think about the difference between syntax rules and morphology rules is that syntax rules, of the form discussed here, are really general formulas for making every single (phrase and) sentence of a language, while morphology rules are individual rules for making specific kinds of words.

### 2.3. How to do syntactic analysis (and find phrase structure rules)

There will be (probably many) times when you’ll be given some sentences in a language that isn’t English, and you’ll be asked to write phrase structure rules that describe the language. Sometimes you’ll also be given some English data that we haven’t studied, and asked to write some new phrase structure rules that describe how the new data works. These phrase structure rule writing tasks are pretty similar. There’s a lot of discussion in this section, but the steps are underlined in the following paragraphs, and summarized at the end of the section.

Start by labelling every word with its part of speech. You’ll need to know all of the words’ parts of speech eventually, because you write rules that say which parts of speech (or which kinds of phrases) go where. Also, if you can look at sentences in terms of the types of words/phrases they contain, rather than literally what words they contain, this can really help with seeing similar structures in different-looking sentences.
Related to this – how to figure out the parts of speech of non-English words:

- If you’re given a non-English sentence with a line under it where every word has an English translation (like below; this line, and the words in it, is called a gloss), every non-English word will have the same part of speech as its English gloss.

- If you’re just given non-English sentences and their English sentence translations, and asked to figure out which non-English word corresponds to which English word, again, once you figure this out, you can basically trust that all non-English words will have the same part of speech as their English translations.

So in this French example below, the parts of speech are as follows; we know this because they’re the same for the French words as for these words’ English correspondents:

<table>
<thead>
<tr>
<th>Det</th>
<th>N</th>
<th>Adj</th>
<th>V</th>
<th>Det</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Une</td>
<td>fille</td>
<td>jolie</td>
<td>mange</td>
<td>le</td>
<td>fromage.</td>
</tr>
<tr>
<td>a</td>
<td>girl</td>
<td>pretty</td>
<td>eats</td>
<td>the</td>
<td>cheese</td>
</tr>
</tbody>
</table>

‘A pretty girl eats the cheese’

Let’s say that the task is to write all the phrase structure rules you possibly can for French based on this data – I’ll ask you to do this pretty often, though I’ll generally give you more than one sentence. But the process is basically the same.

After the words are all labelled with their parts of speech, find all the phrases in the sentence. Another thing you can count on, in foreign data I’ll give you, is that the non-English words which correspond to the words in an English noun phrase also form a noun phrase. So you can find all the phrases in the English translation, then use this information to find all the phrases in the foreign sentence. That is, in the English sentence, a pretty girl is a noun phrase (you should be able to prove this using the pronoun test). This information can be used to find a French noun phrase, which is composed of the French translations of these words: une fille jolie (literally, a girl pretty). Identifying all of the French NPs in this sentence in this manner gives you the following information:

NP
Det N Adj V Det N
Une fille jolie mange le fromage. (because ‘[A pretty girl] eats [the cheese]’)

NP
Det N
Une fille jolie

NP
Det N
le fromage.

NP
Det N
the cheese

Now you have examples of some NPs, and can describe the individual NPs like this:

NP = Det N Adj  (une fille jolie)
NP = Det N  (le fromage)
The information from these schematic structures of individual NPs can be used to write a single rule for the structure of a French NP, which says that a NP consists of a determiner, a noun, and an optional adjective:

$$\text{NP} \rightarrow \text{Det N (Adj)}$$

Next, going back to the goal of writing as many phrase structure rules as possible based on this sentence, look to see what other kinds of phrases you can find in the English sentence. Remember that, as we’ve discussed, every N is part of a NP, every V is part of a VP, and every P is part of a PP. You can use this guideline to make sure you find every single phrase there is to find in the data, starting with English and then finding the corresponding phrases in French. So since you know that eats is a verb, you know there must be a VP around it somewhere; the do-form test can be used to reveal the VP:

$$[A \text{ pretty girl}]_{\text{NP}} \ [eats \ [\text{the cheese}]_{\text{NP}}]_{\text{VP}}$$

This means that the French translations of eats and the and cheese also forms a VP, so we know this:

$$\text{NP} \ \text{NP} \ \text{V} \ \text{NP}$$

This lets us write another rule, this one for the structure of the VP:

$$\text{VP} \rightarrow \text{V NP}$$

One might wonder why this rule isn’t “VP $\rightarrow$ V Det N”. The short answer is that phrase structure rules always refer to the biggest things inside them – so to phrases rather than individual words in rules when possible. Ask me if you’re curious why this is and we haven’t talked about it in class.

Now that we’ve grouped every single word in the sentence into a phrase, we can see and describe the structure of a French sentence:

$$\text{S} \rightarrow \text{NP \ VP}$$

Remember, we generally try to write phrase structure rules for NPs, VPs, PPs, and Ss. So when you’re trying to figure out what’s going on in some new language, you might stop at some point and ask yourself things like, “do I have any evidence for what an NP looks like in this language? Hmm, I should find all the English NPs and then look at their translations.” Same for a VP, and
for a PP, if you find them in the English sentences. Languages will almost always have VPs and NPs, and of course they always have Ss, but you won’t necessarily always get data about PPs.

Finally, a good way to double-check the rules you’ve written is to pick at least two of the sentences you were given and try to construct them using your rules. If you can, you’re in great shape; if you can’t, you know you need to fix something before you turn it in.

Summary: To find phrase structure rules in a foreign language:

1. Label each word in the English translation with each part of speech.

2. Use constituency tests to find all of the phrases in the English translation.

   Every N, V, or P heads a NP, VP, or PP. So to find all phrases:
   • Find each head (N, V, P).
   • Use constituency tests to figure out what NP, VP, or PP that head is a part of.

3. Label each word in the foreign sentence with the part of speech of its English translation.

4. Find the groups of foreign words whose English translations form phrases. Those groups of foreign words also form the same kinds of phrases as their English counterparts.

5. Write one phrase structure rule to describe the structure of each e.g. NP in the data. Do this for all of the phrases: NPs, VPs, PPs, and Ss.

   These rules should describe all of the pieces of the phrase, at the most general level possible. This means describing the contents of a phrase in terms of constituent phrases, instead of individual constituent words, whenever possible.

6. After you’ve done this for each sentence in the data, look at all the rules for a single kind of phrase (e.g. VP) and see if you can combine them at all (see below). Do this for each type of phrase: NP, VP, PP, and S.

You can use basically the same steps to find phrase structure rules in new English data, as follows.

Summary: To find phrase structure rules in English

1. Label all words with their part of speech.

2. Find all phrases in each sentence.

   Every N, V, or P heads a NP, VP, or PP. So to find all phrases:
   • Find each head (N, V, P).
   • Use constituency tests to figure out what NP, VP, or PP that head is a part of.
3. Write individual phrase structure rules to describe the structure of each phrase (each NP, VP, PP, and S).

These rules should describe all of the pieces of the phrase, at the most general level possible. This means describing the contents of a phrase in terms of constituent phrases, instead of individual constituent words, whenever possible.

4. After you’ve done this for each sentence in the data, look at all the rules for a single kind of phrase (e.g. VP) and see if you can combine them at all (see below). Do this for each type of phrase: NP, VP, PP, and S.

On combining sets of phrase structure rules into single rules

Sometimes you can combine rules into a single, more efficient rule; other times you can’t. It’s not strictly wrong to leave rules uncombined, so when in doubt, don’t combine them. But it’s generally better to have one or two general rules than sixteen specific, redundant rules, so when it’s clearly safe, it’ll make me happier (and you might get a couple more points, and also your hand will get less tired) if you combine rules.

A case where you cannot combine rules:

Say you have the following rules:

\[ VP \rightarrow V \ NP \]
\[ VP \rightarrow V \ PP \]

It’s not correct to combine these to form “VP \rightarrow V (NP) (PP)”, because you don’t have evidence that a sentence can have both NP and PP after a V, or that a sentence can have just a V and neither a following NP nor a following PP, or for what order the two would be in if they could cooccur.

A case where you can combine rules:

If you have the following rules:

\[ VP \rightarrow V \]
\[ VP \rightarrow V \ NP \]
\[ VP \rightarrow V \ PP \]
\[ VP \rightarrow V \ PP \ NP \]

…then it’s okay to combine them to “VP \rightarrow V (PP) (NP)”, because you know that a single VP can have both NP and PP (or neither), and you know which order they go in when they combine.
2.4. Drawing syntax trees

Drawing syntax trees is a lot like drawing morphology trees – except that now we’re concerned with the way that words combine into phrases and sentences, and not the way morphemes combine into words. Syntax trees look like this:

```
S
  NP   VP
    Adj  N   V   Adv
      Happy boys  laugh loudly
```

A syntax tree is a representation of the structure of a single sentence – of the way that the individual words are grouped into phrases and eventually form the full sentence. A tree shows every phrase that’s in the sentence, so if your project is to draw a tree, it’s a really excellent idea to start by labelling every word with its part of speech, and then using constituency tests to find all of the phrases in the sentence. In the sentence above, there’s one N and one V, so there’s one NP and one VP; constituency tests tell you what words are in each. You know how to write phrase structure rules for just this sentence based on the constituency tests; they’d look like this:

```
S → NP  VP
NP → Adj  N
VP → V  Adv
```

Very, very importantly, the groupings of words into phrases and into the sentence in a syntax tree absolutely must be an accurate representation of the phrase structure of the language. You’ll usually be asked to draw a syntax tree after you’ve written phrase structure rules for a language. If you just make sure that you’re identifying all phrases when both writing phrase structure rules and drawing trees, and that the phrases in your tree can be generated using your phrase structure rules, this shouldn’t be a problem.

There’s a straightforward relationship between phrase structure rules and trees which you can use to check and make sure you’re consistent: whatever category is to the left of an arrow in a phrase structure rule can, in a tree, be above the categories to the right of the arrow. So we can see a relationship between the rule “NP → Adj  N” and the tree representation of the NP:

```
NP
  Adj   N
    Happy  boys
```

In the tree, the category to the left of the arrow in the rule – NP – is above, and connected to by lines, the categories to the right of the arrow in the rule – Adj and N.
As was mentioned above, because a tree is a representation of a particular sentence, and phrase structure rules ideally describe all possible sentences in a language, a tree has to be consistent with phrase structure rules. This doesn’t mean that a tree has to contain every single constituent listed in a phrase structure rule. To explain this better, let’s look at a second sentence – Boys laugh loudly. This sentence demonstrates that while it’s possible to have an adjective in a NP (like happy boys), it’s not required. We can thus change the NP rule like this:

$$NP \rightarrow (Adj)\ N$$

A tree for the second sentence would look like this:

```
S
   /\  \\
 NP /  VP  \\
   /\       \\
 N V Adv
```

Boys    laugh  loudly

The tree has all the same words in it as the sentence, and it’s consistent with the rule, but because the rule says that adjectives are optional in NPs, and the sentence doesn’t have an adjective in the NP, there’s no reason to draw an Adj in the tree.

**Important properties of trees**

- Every word must be labelled with its part of speech.
- Every phrase present in the sentence must be correctly drawn in the tree.
- Every word must be connected into the tree by a line.
- The groupings of words into phrases and phrases into sentences as shown in the tree must be consistent with the language’s phrase structure rules.

**2.5. Transformations: Syntax rules**

Transformations are ways of changing the word order of a sentence. The fundamental idea behind transformations, and what they’re good for, can be hard to get a grasp on, but once you get a sense for what’s going on, the mechanics of writing and using transformations aren’t that bad. This section will run through the basics of transformations twice, once with lots of words and sentences, and then again more schematically and with more examples.

We assume in linguistics that the order of words in a normal declarative sentence (the kind of sentence you use when you’re observing a fact ‘Rain falls from the sky’ or stating a preference ‘I like cheese’, etc.) is the basic word order of a language. We further assume that any other kind of sentence in which words show up in a different order from these statements is formed by rearranging a set of words that started out in the statement order. In such a rearranged sentence,
we call its original declarative-order structure its **d-structure**, and we call the surface form that’s actually spoken the surface form or **s-structure**.

(Remember that word order is related to tree structure, so we can talk about the order of words in a sentence and also the structure of a sentence interchangeably in this context.)

**d-structure** The original structure of a sentence, when its words are put together in declarative-sentence order. That is, all the words of the sentence, arranged in declarative order.

In declarative sentences, this is the word order that you see on the surface, too.

In sentences where the words are no longer in declarative order, like questions in many languages, this is the structure that the sentence starts out with before a transformation happens to it.

**s-structure** The form of a sentence (i.e. word order and tree structure) that you hear.

In declarative sentences, the s-structure is the same as the d-structure.

In sentences where the words are no longer in declarative order, like questions in many languages, this is the structure that the sentence has after transformations happen to it.

Think about a ‘yes-no’ question (a question that is answered by ‘yes’ or ‘no’), like *Could Bob have some dessert?* A statement that’s very similar to this question – in fact, one you might use to answer it – is *Bob could have some dessert*. Both sentences have exactly the same set of words; the only difference between these two sentences is the word order. So given the assumption that a question like *Could Bob have some dessert?* started out its syntactic life as a statement like *Bob could have some dessert* and then got rearranged, we’re interested in describing the kind of rearrangement that happens to make this sort of question. Under this assumption, the **d-structure** is *Bob could have some dessert*, and the **s-structure** is *Could Bob have some dessert?*

Here, *could* moves from its base position – after the subject of the sentence – to the beginning of the sentence. This movement is called a transformation. A transformation looks a lot like a morphology rule; it has the same four basic parts. This makes some sense – transformations are rules for turning one type of sentence into another, and morphology rules are rules for making one kind of word (usually a root) into another (unlike phrase structure rules, which have a different format and are general formulas for making every single phrase or sentence – well, actually every single d-structure – in a language).
(1) What kind of sentence the transformation produces (some kind of question, etc.).

(2) What change happens to the d-structure to make this new kind of sentence.
   This change is usually moving or adding something; be explicit about what gets
   moved or added.

(3) Where in the sentence this change takes place.
   If something is moved, you need two locations – where something moves from and
   where it moves to.
   If something is added, you only need one location – where it gets added to.

(4) Special conditions, if any, on this change.
   For example, sometimes you do one thing in one situation, and another somewhere
   else – just like allomorphy, where you sometimes add one morpheme and
   sometimes another. If this is the case, make sure to put the conditions for each
   kind of change into the rule.

So a statement of the transformation given above would look like this:

To make a yes-no question, move could from after the subject to before the subject.

(of course, often when you write transformations, you’ll usually move a whole part of
speech, or some type of phrase, or something. I’m just using a simple example where you
move a single word here; the real yes-no question transformation is a bit more complicated)

We can break down the simple transformation above to show how it does these four things:

   (1) To make a yes-no question,
   (2) move could
   (3) from after the subject to before the subject.

Because this is a really simple transformation, there are no special conditions, so no part (4).

Transformations may sometimes involve adding something to a sentence, so it’s not quite true
that d-structures and s-structures will contain exactly the same sets of words. But they will
always come very close to containing the same words. When something is added, it’s a very
consistent thing that gets added to every single e.g. question, and so the addition of whatever can
be described in a rule that can apply to any sentence with any words.

Another, schematic way of saying all this:

ASSUMPTIONS:

Phrases structure rules make all and only strings of words in declarative (statement)
word order.
All other kinds of sentences (like questions, commands, etc.) are made from strings of words in declarative word order, by transformations.

Questions like *Can Bob eat cheese?* and *Could Bob have eaten cheese?* are not in declarative word order; they have an auxiliary before the first NP, unlike statements like *Bob can eat cheese.*

Therefore they must be made by transformations.

Transformations make d-structures into s-structures.

s-structure: the question, sentence, etc. that you say/hear.

d-structure: a string of basically all the same words as the s-structure, but in declarative sentence order.

See the previous section for longer/more precise definitions.

In yes-no questions:

<table>
<thead>
<tr>
<th>d-structure</th>
<th>s-structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob can eat cheese</td>
<td>Can Bob eat cheese?</td>
</tr>
<tr>
<td>Bob could have eaten cheese</td>
<td>Could Bob have eaten cheese?</td>
</tr>
</tbody>
</table>

Note that the d-structure has all the same words as the s-structure, but they’re in declarative-sentence order in the d-structure.

A transformation that makes these d-structures into the corresponding s-structures:

“To make a yes-no question, move the first auxiliary to the beginning of the sentence.”

See the previous section for the parts of a transformation.

More yes-no questions:

<table>
<thead>
<tr>
<th>d-structure</th>
<th>s-structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob eats cheese</td>
<td>Does Bob eat cheese?</td>
</tr>
<tr>
<td>Bob swims</td>
<td>Does Bob swim?</td>
</tr>
<tr>
<td>Bob wants to run a marathon.</td>
<td>Does Bob want to run a marathon?</td>
</tr>
</tbody>
</table>

(Ignore the differences in which verb forms get agreement, i.e. the “s” on the end.)

The d-structure has all of the same words as the s-structure, except for “does”, which is predictably added by the transformation. **The only time when a d-structure should have different words from the s-structure is when the s-structure-only words get regularly added to every sentence added by the transformation.**
The transformation from above, revised to account for this additional data:

To make a yes-no question,
• If there is one or more auxiliary, move the first auxiliary to the beginning of the sentence.
• If there are no auxiliaries, add a form of “do” at the beginning of the sentence.

WH questions: (questions with a WH word: who, what, where, when, why, how)

<table>
<thead>
<tr>
<th>d-structure</th>
<th>s-structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob would eat cheese where</td>
<td>Where would Bob eat cheese?</td>
</tr>
<tr>
<td>Bob would eat what in Paris</td>
<td>What would Bob eat in Paris?</td>
</tr>
</tbody>
</table>

These d-structures aren’t really sentences that you would actually (necessarily) say, but they’re strings of words – the same words as in the s-structures – but in declarative sentence word order.

See the next section for how to find d-structures in this type of situation.

A transformation that makes these d-structures into s-structures:

To make a WH question,
• Move the first auxiliary to the beginning of the sentence.
• Move the WH word from wherever it is in the d-structure to before the moved auxiliary.

This is different, a little bit, from the yes-no questions in that the d-structure isn’t quite a sentence that you would say, and is even less of an answer to the question. That’s fine.

2.6. Figuring out d-structures

If I ask you to find a d-structure, I’ll give you two things: the question that we’re looking for the d-structure for, and a related statement that shows declarative word order. So I could give you this:

What’s the d-structure for: Where would Bob eat cheese?

(If it’s helpful, look at this statement: Bob would eat cheese in Paris.)

The d-structure has to have all the words where, would, Bob, eat, and cheese, but they would have to be in normal declarative order – in the kind of order that the words are in in the declarative sentence Bob would eat cheese in Paris.
The statement tells us where most of the words go – you know you can have the order *Bob would eat cheese*. But that leaves us with the question of where to put *where* in this normal declarative order. *Where* is referring to a location where the cheese-eating is taking place. Locations generally go at the end of the verb phrase – you can see this in the declarative sentence because the cheese-eating location *in Paris* is at the end of the sentence.

So the d-structure is **Bob would eat cheese where**

Once you work out what the d-structure is (and you’ll never have less information than this to use), you can write a transformation that makes it into the s-structure, like above.

It might help to think about it from a speaker’s perspective: If I’m trying to figure out how to say a sentence that asks someone where Bob would eat cheese, I don’t have the words *in Paris* to start with, so I can’t take them out and replace them with *where*. Instead, I’m starting with the word *where*, because I want to know his cheese-eating location. So I put the word *where* in the place in the sentence where cheese-eating locations go in making the d-structure, which my mental grammar will then turn into an s-structure via transformations.
3. **PHONOLOGY**

3.1. The sounds of English, and their features

For the rest of the class, we’ll be talking about phonology, which is concerned with the properties of sounds and the ways that they are combined into words.

**Important:** Sounds, in the sense that we discuss them, are totally different from letters. A word like *through* has seven letters (*t-h-r-o-u-g-h*), but only three sounds (*th-rough*). **DO NOT CONFUSE LETTERS AND SOUNDS.**

To keep letters and sounds straight, we’ll write them differently. When we use words as linguistic examples and just refer to them by their **spelling**, we’ve been italicizing them, like this: *cat.*

To talk about the **sounds** in a word, we do two things differently: we write the words using the International Phonetic Alphabet (IPA), which uses one unique symbol for every sound. This means that sounds like *sh*, that is spelled with two letters, can be written with just one symbol, [*ʃ*]. This avoids confusion – like about when a sequence of letters *sh* stands for one sound, like in *wash*, or for two, like in *misheard*.

The other thing to do when writing **sounds** (rather than letters) is to put them in brackets, like above where I talked about the sound [*ʃ*]. This helps because sometimes the IPA symbols look like regular English letters, and so putting them in brackets makes it very clear that you’re using them as sounds, not letters.

The important parts:
- When **spelling** words, write them like *this* (or *this*), in italics (or underline).
- When writing **sounds**, write them like [*ʃ*], using IPA symbols and brackets.

Here are the symbols to use in transcribing words (i.e. writing the sounds in words):

**IPA symbols for English consonants**

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p</td>
<td>b</td>
<td></td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>v</td>
<td>̃θ</td>
<td>̃ð</td>
<td>̃s</td>
<td>̃z</td>
<td>̃ʃ</td>
</tr>
<tr>
<td>Affricate</td>
<td>m</td>
<td></td>
<td></td>
<td>̃tf</td>
<td>̃dʒ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td></td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td>̃ŋ</td>
</tr>
<tr>
<td>Lateral</td>
<td>m</td>
<td></td>
<td></td>
<td>̃l</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhotic</td>
<td>m</td>
<td></td>
<td></td>
<td>̃r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>̃y</td>
</tr>
</tbody>
</table>

Shaded = voiced
Unshaded = voiceless
Example words

<table>
<thead>
<tr>
<th>Sound</th>
<th>Example</th>
<th>Manner of Articulation</th>
<th>Place of Articulation</th>
<th>Voicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pat</td>
<td>stop</td>
<td>bilabial</td>
<td>voiceless</td>
</tr>
<tr>
<td>b</td>
<td>bat</td>
<td>stop</td>
<td>bilabial</td>
<td>voiceless</td>
</tr>
<tr>
<td>t</td>
<td>pat</td>
<td>stop</td>
<td>alveolar</td>
<td>voiceless</td>
</tr>
<tr>
<td>d</td>
<td>pad</td>
<td>stop</td>
<td>velar</td>
<td>voiceless</td>
</tr>
<tr>
<td>k</td>
<td>kat</td>
<td>stop</td>
<td>velar</td>
<td>voiceless</td>
</tr>
<tr>
<td>g</td>
<td>get</td>
<td>stop</td>
<td>glottal</td>
<td>voiceless</td>
</tr>
<tr>
<td>f</td>
<td>fat</td>
<td>fricative</td>
<td>labiodental</td>
<td>voiceless</td>
</tr>
<tr>
<td>v</td>
<td>vat</td>
<td>fricative</td>
<td>labiodental</td>
<td>voiceless</td>
</tr>
</tbody>
</table>

How do I know what features a sound has, or what symbol to use for it?

The easiest way is to look at the lists of example words for each sound and find a sound (and symbol) that’s the same as the one you’re trying to identify. You can then use that symbol in a transcription, or look the sound/symbol up on the IPA chart and read off its features.

You can also get a deeper understanding of what these features mean by making the sound, figuring out what your head is doing when you make it, and use that to identify the features that the sound has, as follows.

Each consonant has three features:

- Manner of articulation (stop, fricative, affricate, nasal, lateral, rhotic, or glide)
- Place of articulation (bilabial, labiodental, interdental, alveolar, palatal, velar, or glottal)
- Voicing (voiced or voiceless)

**Manner of articulation (what kind of thing your head is doing)**

Stop: No air comes out of the mouth (or nose) at first, because the tongue or lips are fully closed and block it; then the tongue or lips open up and air comes out (sometimes in a strong burst).

Test: Put your mouth in position to start saying the sound. Try to exhale. If you can’t make air leave your lungs because your vocal tract is completely closed, it’s a stop.

Fricative: The tongue or lips come very close to each other (or to the teeth, or roof of the mouth) and make a very small opening, which air hisses through, making a friction-like sound.

Test: Make the sound for a while, and put your hand in front of your mouth. If you hear continuous hissing, and/or can feel air quickly leaving your mouth, it’s a fricative.
Affricate: The mouth is closed at first, as in a stop; then it opens a little bit into position for a fricative.

Test: If the sound has two parts, and the first part passes the stop test and the second part passes the fricative test, it’s an affricate.

Nasal: No air comes out the mouth; instead, air comes out through the nose.

Test: Rest your fingers on the bridge of your nose and make the sound; if your nose vibrates, it’s a nasal. Also, hold your nose and make the sound. If your mouth bizarrely fills up with air and/or you can’t hold the sound for long, it’s a nasal.

Lateral, rhotic, glide: These are all articulations where your tongue makes little or no contact with the roof of your mouth, allowing lots of air to flow out of the mouth.

Test: If a sound isn’t any of the above, it’s one of these. That’s the best I can do.

Voicing

Voiced: The vocal cords vibrate when the sound is made.

Voiceless: The vocal cords do not vibrate when the sound is made.

Test: Put your hand on your throat, and hold the sound for a while. If you feel your throat (actually, your larynx) vibrating, the sound is voiced. If you don’t, the sound is voiceless.
Place of articulation

Make the sound, and hold it for a while. Feel where your tongue (or your lips) is touching some part of your mouth. Then find that part of your mouth on the picture below.

**FIGURE 1.7** A sagittal section of the vocal tract, showing the places of articulation that occur in English. The coronal region is shown in more detail at the right.

(don’t worry about the retroflex or palato-alveolar places; we don’t use them in English)
IPA symbols for English vowels

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Mid</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td><strong>Mid</strong></td>
<td>i</td>
<td>ø</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>ø</td>
<td>o</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>æ</td>
<td></td>
<td>a</td>
</tr>
</tbody>
</table>

Shaded = rounded
Unshaded = unrounded

Example words

\[\begin{array}{l|l|l}
  i & beet & [bit] \\
  ɪ & bit & [bit] \\
  ë & bait & [bet] \\
  e & bet & [bet] \\
  æ & bat & [bæt] \\
\end{array}\]

\[\begin{array}{l|l|l}
  u & boot & [but] \\
  ʊ & book & [buk] \\
  ø & but & [bat] \\
  ø & bought & [bot] \\
  a & father & [faðər] \\
\end{array}\]

Each vowel has three features:
- Height (high, mid, or low)
- Backness (front, mid, or back)
  (note that “mid height” is a different “mid” than “mid back”)
- Rounding

3.2. Minimal pairs, and an introduction to phonemes vs. allophones

A minimal pair is a pair of words that differ in the smallest possible way – that is, a pair of words that differ minimally. The smallest unit of sound is a phoneme, so a minimal pair is a pair of words that differ in a single phoneme (basically, in a single important sound).

Phoneme A distinctive sound in a language. Changing one phoneme in a word to another makes a new word in the language.

Minimal pair A pair of words that differ only by a single phoneme.

Examples \( [pæt] \sim [bæt] = pat \sim bat \)
\( [pæt] \sim [pæt] = pat \sim pit \)
\( [pæt] \sim [pæd] = pat \sim pad \)

Minimal pairs are useful because they let us identify phonemes. That is, because changing \([p]\) to \([b]\) at the beginning of an English word makes a whole different word, the difference between \([p]\) and \([b]\) is phonemic in English.
It’s important to know that you can also pronounce a sound differently without making a whole new word. For example, there are two ways of pronouncing the sound we think of as p in English: the ps in *apple* and *appall* are fairly different, in that the second p is followed by a puff of air that you can feel by putting your hand in front of your mouth when you pronounce the words. (That p with the puff of air is also the one where you might spit on someone.) This difference is represented by transcribing the p in *apple* as usual – [p] – but transcribing the puff of air p in *appall* with a little h floating after it, like this [pʰ]. That puff of air is called ‘aspiration’, and so [pʰ] is called ‘aspirated p’. Now, if you were to pronounce *apple* like [æpʰəl], with the aspirated p, native English speakers should agree that that could never be a different English word from the normal pronunciation [æpəl], even though the two do sound different. If the word were pronounced as [æbəl], though, with a [b] instead of a [p], that could be a different word (even though it’s not a word we know).

What this shows is that the difference between [p] and [b] is phonemic (that is, meaningful) in English, while the difference between [p] and [pʰ] is not meaningful – those just feel like two different ways of pronouncing the same sound. In the next section, we’ll discuss those non-meaningful differences, which are called allophonic differences.

One more important thing to point out is that not all languages consider the same differences meaningful. We can see a simple example of this by continuing to think about the sounds [pʰ], [p], and [b]. In English, there are two meaningful categories made up of these sounds, like this:

| English | pʰ | p | b |

Other languages also use these three sounds, but categorize them differently. For example, in languages from India like Hindi, these three sounds are all meaningfully different; that is, there are three different words [pʰal] ‘knife blade’, [pal] ‘take care of’, and [bal] ‘hair’. This shows that there are three different meaningful categories made up of these sounds in Hindi, like this:

| Hindi | pʰ | p | b |

In languages from Australia like Warlpiri, on the other hand, there is only one category made up of all three of these sounds. That is, there are no minimal pairs in which two words differ only in that one has one of these sounds and the other has another. Instead, just like [pʰ] and [p] sound like ‘different ways of saying one sound’ to an English speaker, all three of these sounds sound like different ways of saying a single sound to a Warlpiri speaker, like this:

| English | pʰ | p | b |

What this shows is that while lots of languages can use the same sounds, the languages’ grammars can categorize these sounds differently. Therefore, to know about the sound system of a language, we not only need to know what sounds someone pronounces when they speak that language, but also need to know about things like minimal pairs to understand how languages categorize – and so how they’re really using – the sounds.
3.3. Allophones and complementary distribution

Sometimes a single phoneme can be pronounced more than one way. This happens a lot in English – for example, say the words *pat* [pæt] and *pan* [pæn]. Both words have the same vowel phoneme, but the vowels sound different. This is because the nasal [n] at the end of *pan* makes the vowel sound nasal, too – you should be able to feel your nose vibrating during the vowel in *pan*. We can note this explicitly in the transcriptions by writing [pæn], where the [ ~ ] indicates that the vowel is nasal.

We know that [æ] and [ã] aren’t different phonemes in English, for two reasons. First, ‘mispronouncing’ [æ] as [ã] doesn’t make a new English word – if you heard *pat* mispronounced as [pæt], where it contained the *pan* vowel rather than the normal *pat* vowel, it would sound wrong, but not like a word other than the intended *pat*. (To me, actually, you’d sound like a cool girl from Michigan.) Second, there are no minimal pairs that differ only in that one word has [æ] where the other has [ã]. When a single phoneme can be pronounced in two or more ways, these possible pronunciations are called allophones. Remember that allomorphs are multiple different ways of pronouncing a single morpheme; similarly, allophones are various pronunciations of a single phoneme.

**Allophones** Multiple pronunciations of a single phoneme; the choice of which allophone to use in which context is predictable.

We’ll be doing a lot of looking at sets of words, in English or some other language, and trying to figure out whether pairs of sounds are independent phonemes or allophones of a single phoneme – that is, whether the difference between a given pair of sounds is phonemic or allophonic.

The best and easiest way to discover/prove that the two sounds are phonemes is to find a minimal pair in the data you’re given.

If there are no minimal pairs proving that the difference between the sounds is phonemic, you can instead look for the following kind of evidence that the sounds are allophones. Allophones are predictable – that is, when a phoneme can be pronounced in more than one way, these two pronunciations aren’t used randomly, wherever a speaker feels like it – instead, they’re predictable from the context in which the sound is appearing.

To see this more clearly, here are some words where [æ] and [ã] appear:

- *pad* [pæd]
- *man* [mæn]
- *hat* [hæt]
- *Sam* [sæm]
- *wrath* [ræθ]
- *hang* [hæŋ]
- *mass* [mæs]

If you weren’t an English speaker, and this data were all you had to decide whether [æ] and [ã] were allophones, the best way to do this is to figure out whether the two sounds are in complementary distribution; that is, whether [æ] always shows up in one context in words, while
[æ] always shows up in a different context in words. Since the pronunciation of allophones is predictable, finding this sort of regular, predictable pattern indicates that two sounds are in fact allophones.

**Complementary distribution** Allophones are in complementary distribution when they appear predictably in different contexts – that is, when the distribution of one allophone complements (i.e. is strictly different from) the distribution of the other.

Going back to the words above, we should look at the contexts where [æ] and [æ] appear and see if there’s a way to predict which of those sounds gets used in which word. When you’re looking at sounds’ contexts, it’s usually easiest to start by looking at the segments before and after the sounds and see whether they have anything in common. An easy way to do this is to make lists of the sounds’ preceding and following segments, and then see if the collections of contexts have anything in common.

| Before [æ] | [p], [h], [r], [m] |
| Before [æ] | [m], [s], [h] |
| After [æ]   | [d], [t], [θ], [s] |
| After [æ]   | [n], [m], [η] |

The sounds before [æ]/[æ] don’t predict which sound to use – both [æ] and [æ] can follow [m] and [h]. Therefore, you can’t use the existence of a preceding e.g. [m] to tell you whether you should use the allophone [æ] or [æ] in a given context – either allophone could appear there.

The two vowels are followed by totally different sets of sounds, though, so one could use the sound following the vowel to figure out which pronunciation to use. Looking more carefully, the sounds that follow [æ] are all nasals, and the sounds that follow [æ] are all not nasals. We can thus predict which sound appears where, like this:

[æ] always appears before non-nasal segments; [æ] always appears before nasal segments.

Sometimes the context determining the pronunciation of an allophone might be more complicated – it could be the preceding or following vowel or consonant, rather than simply the preceding or following segment. It could also have something to do with a segment’s position at the beginning or end of a word, or in a stressed or unstressed syllable. Sometimes it might be the preceding and following segments. And sometimes, there might be more than one feature on a neighboring segment that determines an allophone’s pronunciation. But this method of listing environments and looking for similarities should help you find all of these.
Possible contexts that can determine the pronunciation of an allophone

- Preceding segment
- Following segment
- Preceding/following vowel
- Preceding/following consonant
- Beginning of a word
- End of a word
- In a stressed syllable
- In an unstressed syllable
- etc.

It’s really, really important, when you’re listing the things that can come before or after the sounds you’re looking at, to remember to list word edges. Sometimes one allophone might appear at the beginning of a word and the other might appear everywhere else; if you don’t somehow write down “beginning of a word” (I like to write “∅”, for “no sound”, for short) in the “things coming before the sound” list, you could easily miss this generalization.

Are [X] and [Y] phonemes or allophones? A summary of how to work it out:

Is there a minimal pair of words, differing only in that one has [X] where the other has [Y]?

YES  \( \rightarrow \) [X] and [Y] are phonemes.

NO  \( \rightarrow \) no information.

Look at the places where [X] and [Y] appear in words (their contexts).
Are [X] and [Y] in complementary distribution? That is, does either sound have a consistent, predictable context, such that you could say something like:
“[X] is always before sounds with feature Z; [Y] is never before sounds with feature Z.”?

YES  \( \rightarrow \) [X] and [Y] are allophones.

NO  \( \rightarrow \) [X] and [Y] are phonemes.

A PHONEMIC difference between two sounds is a meaningful difference – trading one sound for the other in a word can make a new word.

An ALLOPHONIC difference between two sounds is a predictable difference – specific properties of the context of the sound tells you which of the two should appear there.

3.4. Finding the basic phoneme, and phonetic motivation for allophonic changes

The last section talked about two segments (like [æ] and [ʌ]) being allophones in a general way. We can get a little more specific and say that one sound is an allophone of the other, meaning that there’s one fundamental sound, which sometimes gets pronounced as itself and other times turns into the other sound in particular contexts.

In the example above, we could say that [ʌ] is an allophone of [æ], meaning that [æ] is the basic phoneme, and that it turns into [ʌ] in a particular context – that is, before a nasal consonant.
There are a couple reasons for doing this. First, it allows us to write simple, schematic rules to describe the distribution of allophones, as will be described in the next section. Second, it allows us to capture the fact that allophonic variation is generally phonetically motivated. That is, most of the time, when two sounds are allophones and so have predictable distributions, the variation between the two sounds is such that sounds appear in contexts that they have something in common with.

In the [æ}/[ə] example, [ə] appears before nasals, and it is itself a nasal vowel. You can tell that [ə] physically has something in common with nasals because just like your nose vibrates when you say [m] or [n], it also vibrates when you say [ə]. In this way, if we say that [æ] turns into [ə] before nasals, we’re really saying that when [æ] is followed by a nasal, it becomes like that following nasal by turning into [ə]. This is, in the simplest terms, a way in which your head makes speaking easier for itself.

Finding the basic phoneme can be a little tricky, but here’s the general rule of thumb:

If [X] and [Y] are allophones, and

- [X] appears in a very limited context (like only before nasals), while
- [Y] appears in a more general context (like before every other kind of sound, and at the end of words)

\[ \rightarrow \text{...then the sound with the general distribution (here, [Y]) is the basic phoneme, and} \]
\[ \rightarrow \text{...the sound with the limited distribution (here, [X]) is an allophone of the other.} \]

(\[ \rightarrow [X] \text{is an allophone of [Y].} \])

That is, it’s generally true that the basic phoneme is the one that shows up all over the place, and it turns into an allophone in a very limited set of contexts.

3.5. Phonological rules

When we talked about words turning into other kinds of words via morphology, or sentences turning into other kinds of sentences via transformation, we wrote rules to describe these changes. We’ll also write rules for phonological changes – here, phonemes turning into allophones. These rules, like morphological and syntactic rules, are sentences with defined parts.

The parts of a phonology rule are as follows:

1. What the basic phoneme is
2. What allophone it sometimes turns into
3. Where this change occurs (what the phonological context of the allophone is).
In the [æ]/[æ] example that we’ve been talking about, we’ve worked out that the phoneme [æ] turns into its allophone [æ] before nasals. Here’s what the formal rule saying this looks like – it’s an awful lot like that sentence, actually:

(1) The basic phoneme [æ]
(2) becomes the allophone [æ] before a nasal consonant.

One important thing to note is that a rule can be about a single phoneme turning into an allophone in a particular context, but it can also be more general – in English, it’s actually true that all vowels become nasalized before nasal consonants. So you could write a more general rule describing this fact like this:

A basic phoneme, which is a non-nasal vowel, becomes an allophone, which is the same vowel but nasalized, before a nasal consonant.

Whatever the properties that distinguish a set of basic phonemes you might want to talk about are, just list them in the first part of the rule. Whatever features they pick up or changes they undergo when becoming their allophones, explain them in the second part of the rule. And whatever features of the context define the change get listed in the third part of the rule.

One comment about rule-writing and finding basic phoneme forms from sets of allophones – rules show how to turn a basic phoneme into an allophone, so you generally need to choose a basic phoneme before you can write a rule. But you don’t actually have to choose the correct basic phoneme to write a rule that gets the job done. For example, back to the [æ]/[æ] example – let’s say we’d (wrongly) decided that [æ] was the basic phoneme, and that [æ] was an allophone of [æ]. In this case, while we’ve chosen the wrong basic phoneme, we can still write a correct rule that describes [æ] turning into [æ] and gets the overall distribution of [æ] and [æ] right, like this:

The basic phoneme [æ] becomes the allophone [æ] when it’s not before a nasal consonant.

If this were a homework or exam problem, this answer would miss the “find the basic phoneme” points, but would get the “write a rule that does the right thing” points.

3.6. Positions in syllables

There’s a cheesy saying in linguistics that “syllables are what syllable has three of.” Syllables are pretty hard to define without some formal terms, which I’ll describe in a minute. But if you’re trying to figure out how many syllables there are in a word, you could try clapping or tapping your foot or something while you say a word slowly – most people find that they’re able to clap once per syllable. Try it with big long words like de.reg.u.lation and an.ti.ma.cass.a.rite and de.pro.nom.i.na.li.zation (which are divided into syllables with periods).
Once you have a sense for what syllables are, we can start talking about what parts they have. Words are, of course, made up of vowels and consonants, and syllables are (also of course) arrangements of vowels and consonants. The first thing to know is that all syllables (well, almost all – can you think of counterexamples in English?) have at least one vowel in the middle. This vowel (or these vowels) is (are) called the nucleus. In a very simple word like cat [kæt], which consists of a single syllable, the nucleus is the vowel [æ].

If there are any consonants before the nucleus, these consonants are the onset of the syllable. If there are any consonants after the nucleus, these consonants are the coda of the syllable. Finally, sometimes we might want to talk about the nucleus and coda together – this part of the syllable is called the rime (because words that rhyme have to have the same final nucleus and coda – like in cat – pat [kæt] – [pæt]).

If a word has more than one syllable, we indicate syllable boundaries using a period, like this: syll.able [sɪl.ə.bʌl].

**Nucleus**  
The vowel(s) in a syllable.  
All syllables have nuclei.

**Examples**  
cat [kæt], at [æt], see [si], eye [ai], craft [kræft], syll.able [sɪl.ə.bʌl]

**Onset**  
Any consonants in a syllable that come before the nucleus.  
Not all syllables have onsets.

**Examples**  
cat [kæt], see [si], craft [kræft], syll.able [sɪl.ə.bʌl]  
Syllables without onsets  
at [æt], eye [ai]

**Coda**  
Any consonants in a syllable that come after the nucleus.  
Not all syllables have codas.

**Examples**  
cat [kæt], at [æt], craft [kræft], syll.able [sɪl.ə.bʌl]  
Syllables without codas  
see [si], eye [ai]

**Rime**  
The nucleus and coda of a syllable.  
Because all syllables have nuclei, all syllables have rimes.

**Examples**  
cat [kæt], at [æt], see [si], eye [ai], craft [kræft], syll.able [sɪl.ə.bʌl]

Finally, sometimes we will talk about complex syllable constituents. A nucleus, onset, or coda is complex if it contains more than one phoneme. So in the example words above, eye has a complex nucleus, because there are two vowels in the nucleus [ai]; craft has a complex onset because there are two consonants in the onset [kræft] and also a complex coda because there are two consonants in the coda [kræft].