I-Language

An Introduction to Linguistics as Cognitive Science

Daniela Isac
and Charles Reiss

OXFORD
UNIVERSITY PRESS

they occur in exactly the same environment, as in *wetting* and *wedding*. Most speakers of dialects in which these words are complete homophones will insist that the words are pronounced differently. They are not, unless one produces a spelling pronunciation, a pronunciation that does not reflect the grammar, as in *I said* weld[ɪ]ng, *not* welt[ɪ]ng, but this does not reflect what the grammar does with /t/ and /d/ inputs in this environment.

We should note that it is perfectly possible to be a native speaker of English and treat flap and plain [t] as allophones of a single phoneme, and also be a native speaker of Japanese, in which [t] and flap represent completely distinct phonemes. Or to be a speaker of both English and Thai, and thus treat plain and aspirated *t* as allophones in English but as separate phonemes that can distinguish word meaning in Thai. In other words, a bilingual person will process information and thus construct experience differently depending on which of the two I-languages is used. The equivalence class is not in the signal but is a result of signal processing.

Before moving on, we reiterate that your own dialect of English may be completely different from ours. For example, you may not have any flaps at all; and you may have glottal stops in places where we do not, as in words like *writer*, as opposed to a [d] in *rider*. We have a flap in the middle of both of these words, and they only differ in terms of the vowels preceding the flap. This is typical of North American English, but it may be hard to convince speakers that it is true. We have provided a recording on the companion website.

6.4 Turkish vowel harmony

We are now ready to return to the analysis of Turkish vowel harmony, the phenomenon mentioned at the very beginning of the book. Consider the data in (6.17) that shows the nominative singular, nominative plural, the genitive singular and the genitive plural of eight Turkish nouns. Nominative is the form used when the noun is subject of a sentence, so *ipler* would be the word for ropes in a sentence meaning "The ropes fell." The genitive is used to show possession, so *ipin* means "the rope's, of the rope." The genitive plural combines the meanings genitive and plural, so *iplerin* means "of the ropes."

6.17

Turkish vowel harmony data [18]

	nom. sg.	nom. pl.	gen. sg.	gen. pl.	
a.	ip	ip-ler	ip-in	ip-ler-in	"rope"
b.	kıl	kıl-lar	kıl-ın	kıl-lar-ın	"body hair"
c.	sap	sap-lar	sap-ın	sap-lar-ın	"stalk"
d.	uç	uç-lar	uç-un	uç-lar-ın	"tip"
e.	son	son-lar	son-un	son-lar-ın	"end"
f.	öç	öç-ler	öç-ün	öç-ler-in	"revenge"
g.	gül	gül-ler	gül-ün	gül-ler-in	"rose"
h.	ek	ek-ler	ek-in	ek-ler-in	"joint"

As you look over the data you will notice that the plural suffix takes two different forms *-ler* and *-lar*. In the genitive singular column, you notice four different forms: *-in, -ın, -un, -ün*. In the genitive plural column we see the same two forms of the plural, *-ler, -lar*, but only two forms of the genitive marker.

Here are some questions about these forms:

• What determines the choice of vowel in each suffix?
• How can we represent the suffix?
• Do we have to say that the genitive suffix in the plural is different than the genitive suffix in the singular, since the former has only two forms and the latter has four?

In order to begin to answer these questions, we need to understand a bit about the phonetic correlates of Turkish vowels. We will need to describe them along three dimensions that we will illustrate using English vowels. First, pronounce the word *beat* slowly to yourself and try to pay attention to the position of your tongue and jaw. For our purposes, what is important is just to compare this vowel with that of *bet*. In the latter, you should feel (or see in a mirror) that your tongue and jaw are lower than in the former. We will refer to the vowel of *beat*, whose phonetic symbol is [i], as a HIGH vowel, and to the vowel of *bet*, as a NON-HIGH vowel. For this vowel, we will use the symbol [e], since we will follow the Turkish writing system in this discussion, although the symbol used in the International Phonetic Alphabet is [ɛ].

Next, compare, by feeling and looking in a mirror, the vowel of *beat* with that of *boot*, which we will write as [u]. You should notice that in *beat* your

---

18 The symbol ç represents the sound written *ch* in English. The vowel symbols will be explained in the main text.

lips are spread rather wide, whereas in *boot* the lips are rounded. The vowel in *boot* is ROUND, that in *beat* is NON-ROUND.

The last parameter is a bit harder to notice, but it can also be illustrated using *beat* and *boot*. Try to say the words slowly and silently, concentrating not on the lips but on the difference in the position of the tongue. One way to isolate what is happening with the tongue is to try to say a long version of the vowel in *boot* but forcing the lips to stay spread as for the vowel in *beat*. You won't sound natural, and you will look funny, but you should notice that the tongue is pulled further back in the mouth for the *boot* vowel than for the *beat* vowel. The latter is called a NON-BACK vowel and the former a BACK vowel.

The three binary choices NON-BACK vs. BACK, and so on, allow for 2×2×2 combinations, each one corresponding to one of the eight Turkish vowels in the roots of the words in (6.17), as shown in the following table:

	NON-BACK		BACK	
	NON-ROUND	ROUND	NON-ROUND	ROUND
HIGH	i	ü	ı	u
NON-HIGH	e	ö	a	o

You can play with your speech articulators and figure out the approximate pronunciation of the other vowels in the table. For example, [ü] has the tongue forward and high like [i], but it has the lips rounded, like [u]. This sound occurs in German, also written *ü*, and in French, where it is written *u*. If you are familiar with the International Phonetic Alphabet, you will know that the symbol for this vowel is [y].

The photographs in Fig. 6.2 will give you an idea of how these vowels are articulated. You can hear sound files and see more photos on the companion website.

The descriptions in terms of the configurations of the vocal tract that we have provided correspond to equivalence classes of vowels. The label of each class (for example HIGH) is typically called a *distinctive feature* or just *feature* in the linguistics literature. The features correspond to the linguistically relevant distinctions among speech sounds in the languages of the world. Sets of segments that correspond to a feature description are called *natural classes*. For example, the set of HIGH vowels is a natural class (containing i, ü, ı, u), as is the set of HIGH, NON-ROUND vowels (containing i, ı). Note that the more features we list, the smaller the natural class is, because adding features makes the description more specific.
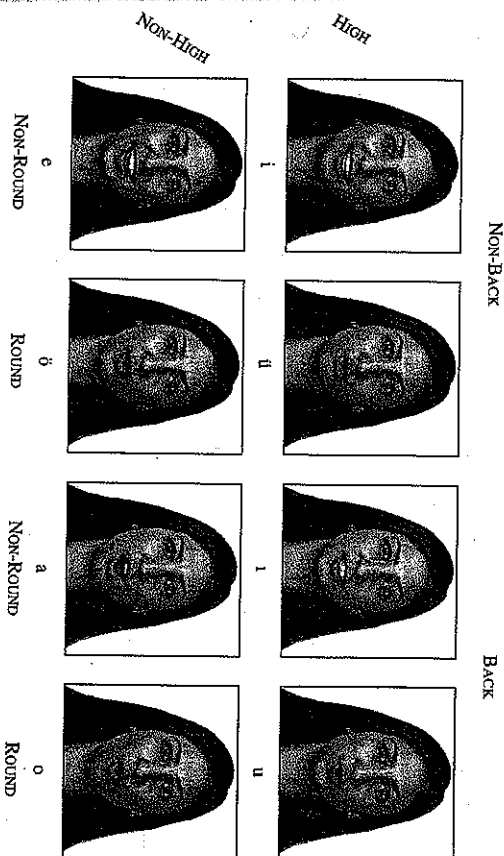
Fig 6.2 A native speaker pronouncing the eight Turkish vowels. See companion website for side views and sound files. The photographer, Sabina Matyiku, and the model, Ezgi Özdemir, are both Concordia undergraduate linguistics students.

For practice, list the members of the natural class of ROUND vowels in the table. Now list the NON-HIGH, ROUND vowels. Now list the BACK vowels.

So we are now ready to analyze the Turkish forms—we have broken down the data to make the presentation clearer. We begin with just some of the nominative singular and plural forms you saw above, as well as some new forms that will make the patterns more apparent. Try to answer the questions before you read the discussion that follows.

6.18   Turkish singular/plural pairs

singular	plural	meaning
dev	devler	"giant"
kek	kekler	"cake"
can	canlar	"soul"
cep	cepler	"pocket"
tarz	tarzlar	"type"
kap	kaplar	"recipient"
çek	çekler	"check"
saç	saçlar	"hair"

şey   şeyler   "thing"
ters   tersler "contrary"
aşk   aşklar   "love"

a. What are the two forms of the plural suffix? 1.

   2.

b. What determines where you find each suffix?
   • Suffix (1.) occurs ...
   • Suffix (2.) occurs ...

Notice that the plural of *sap* is *saplar* and the plural of *ek* is *ekler*. It turns out that this is a general pattern—if the root has just an *e*, then the plural is -*lar*; if the root has just an *a*, then the plural is -*ler*.

What about the other roots that take -*lar* in the plural? They are *uç, kıl, son*. What do you notice about the four vowels *a, o, u, ı?* ... Right, they are all BACK. And the four vowels that take the suffix -*ler* are all NON-BACK: *e, ö, i, ü.*

6.19   More Turkish singular/plural pairs

singular	plural	meaning
ip	ipler	"rope"
kıl	kıllar	"body hair"
sap	saplar	"stalk"
uç	uçlar	"edge"
son	sonlar	"end"
öç	öçler	"vengeance"
gül	güller	"rose"
ek	ekler	"junction"

a. What are the two forms of the plural suffix? 1.   2.

b. What determines where you find each suffix?
   • Suffix (1.) occurs ...
   • Suffix (2.) occurs ...

So to compute which vowel occurs in the plural, we read the NON-BACK/BACK value of the vowel on the bare singular root form and choose the version of the plural whose vowel value agrees with it.

Let's be explicit—what does this show us about equivalence classes? Well, the vowels that are, say, NON-BACK, are not all pronounced with the tongue in exactly the same place—for the purposes of phonology, lots of physical detail is ignored, and these vowels can be treated as identical in some respect.

Why does this illustrate computation? Well, we can formulate an explicit algorithm referring to symbols that corresponds to the patterns we see in the data. Something like this algorithm seems to underlie the behavior of Turkish speakers.

Before we proceed, we should note that Turkish speakers are not physically constrained to obey the patterns of vowel harmony. For instance, they may be bilingual in Turkish and English, so as people their behavior is not always vowel-harmonic. Also, we can ask a Turkish speaker to pronounce, say, *sap-ler*, and he or she will be able to do so—but this does not reflect the computations of the grammar. Grammars do not directly model behavior, and behavior is just one of the sources of evidence for understanding the computations of the grammar.

So what *is* the plural suffix in Turkish? How is it stored in the mind of a speaker? Is its vowel encoded as NON-BACK or BACK? Let's ask some easier questions. What does the plural suffix start with? It seems reasonable to assume that it starts with an *l*, since there is no evidence that it is anything else. We won't worry here about how to express *l* using features (we would need many new ones). It also seems reasonable that it ends with *r*. So the form is something like -*lVr*, where V is the vowel we need to figure out.

The V in the middle of the plural suffix appears never in a rounded form, and it appears never as a HIGH vowel, so it seems reasonable to assume that the stored form is a member of the equivalence classes NON-HIGH and NON-ROUND. What about the value for NON-BACK/BACK? Well, we could assume that it is NON-BACK, basically that the vowel is stored in memory as *e*, and then change it into *a* when it follows a BACK vowel. But we could also assume it is basically BACK and have a rule that changes it to *e* when it follows a NON-BACK vowel.

These two options seem equally valid, and there is no principled way to choose among them. If all we wanted to do was write a computer program to mimic Turkish output, then it would not matter which one we use. However, the cognitive biolinguistic approach assumes that there is some truth to the matter concerning what computations characterize a Turkish grammar.[19] We have a third option, in addition to choosing between *a* and *e*.

19 Obviously, different speakers could have different versions of the rule, even if their grammatical output is the same.

First, let's recognize that a child learning Turkish would find him- or herself in a situation like that of us, the linguists—there is no principled way to decide between basic e or a. If part of the goal of grammatical theory is to explain how children learn language, then leaving them to make random choices is not much of a solution. A second point to appreciate is that, once we have features, the symbols e, a, and so on become unnecessary. These symbols are just abbreviations for highly specific equivalence classes: e is just an abbreviation for the intersection of the sets denoted by NON-BACK, NON-HIGH, and NON-ROUND. This realization liberates us to propose that the equivalence class in memory that corresponds to the vowel in the plural suffix is just characterized as NON-HIGH, NON-ROUND—with no value for the third contrast NON-BACK/BACK. In other words, the vowel in the suffix has no value along this dimension and the computational system, the phonological grammar, provides it with one. The stored form of the vowel is thus something like this:

$$\begin{bmatrix} V \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix}$$

The plural suffix then contains this vowel preceded by an l and followed by r. As noted above, we will not provide featural representations for l and r, and thus we represent the suffix thus:

$$l \begin{bmatrix} V \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix} r$$

This conclusion not only frees us from being forced to make an arbitrary choice but also leads to an elegant analysis of the rest of the data, as we will now see.

We turn now to an analysis of the genitive singular forms. We need to first identify what forms occur, then determine the environment in which each variant occurs, and finally posit a form that is stored in memory that can be used to compute the output forms.

6.20 Turkish nominative and genitive singular pairs

nom. singular	genitive singular	meaning
ip	ipin	"rope"
kıl	kılın	"body hair"

sap	sapın	"stalk"
uç	uçun	"edge"
son	sonun	"end"
öç	öçün	"vengeance"
gül	gülün	"rose"
ek	ekin	"junction"

a. What are the four forms of the genitive suffix?

b. What determines where you find each suffix?

1.   2.   3.   4.

- Suffix (1) occurs ...
- Suffix (2) occurs ...
- Suffix (3) occurs ...
- Suffix (4) occurs ...

Here are the four vowels that occur in the suffix: i, ü, ı, u. What do they have in common? They are all HIGH, so it seems reasonable to suppose that this suffix is stored in memory with a vowel encoded as HIGH.

Now, where does the form -in occur? It occurs in the forms ipin and ekin. These are the forms with root vowels that are NON-BACK, NON-ROUND, which agrees with the vowel i of the suffix. Where does the suffix -ün occur? It occurs in the forms öçün and gülün, which have vowels that agree with the suffix vowel in being NON-BACK, ROUND. We get the same pattern with the other two forms of the suffix: -ın occurs with sapın and kılın, which both have BACK, NON-ROUND vowels; and -un occurs in uçun and sonun, where the vowels are BACK, ROUND.

We now see that we don't have to choose a particular vowel to represent the genitive suffix. It is stored in memory as containing a vowel that is specified as just HIGH, and that's it:

$$\begin{bmatrix} V \\ \text{HIGH} \end{bmatrix} n$$

The other features are filled in by the computational system to match those of the vowel that precedes.

The genitive suffix is thus encoded in memory as an abstract entity that starts with a vowel specified as HIGH followed by an n. The other features of the vowel get filled in depending on the context by the grammar. Note that the stored forms of the vowels of the plural and the genitive are both lacking specification for some features, but the two vowels are distinct from each other.

We are now ready to tackle the genitive plurals.

6.21   Turkish nominative singular/genitive singular/genitive plural pairs

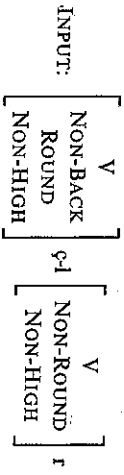| nom. singular	genitive plural		meaning
ip	iplerin		"rope"
kıl	kıllarin		"body hair"
sap	saplarin		"stalk"
uç	uçların		"edge"
son	sonların		"end"
öç	öçlerin		"vengeance"
gül	güllerin		"rose"
ek	eklerin		"junction"

a. What are the two forms of the genitive suffix in this data?
   1.
   2.
b. What determines where you find each suffix?
   • Suffix (1.) occurs ...
   • Suffix (2.) occurs ...

6.22   Input to the phonology for öçler

$$\text{INPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}\ \text{ç-l}\ \begin{bmatrix} V \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix}_r$$

Recall that we wondered if the genitive marker in the genitive plural needed to be encoded differently from that in the singular, since the former shows up in only two forms, whereas the latter has the four combinations of NON-BACK/BACK and ROUND/NON-ROUND discussed above. The genitive marker in the plural has only the NON-ROUND variants. What do you think—are there two genitive markers, one for singular and one for plural, or can the two patterns, a four-way contrast and a two-way contrast, all be derived from a single underlying form?

We propose that there is a single genitive suffix underlyingly. To see how this works, consider the system of morphology and phonology that we are proposing. The morphology takes roots and suffixes as they are stored in memory and puts them together. That will give us structures like the following for a form that exits the grammar as öçler, with the vowels expressed in terms of features. We use phonetic symbols for the consonants for ease of presentation.
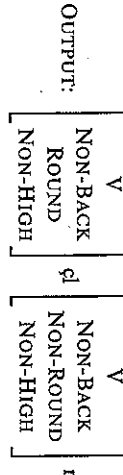
Since the suffix vowel is lacking a specification for NON-BACK/BACK, it looks to its left and copies the first one it finds—in this case the NON-BACK of the vowel ö.
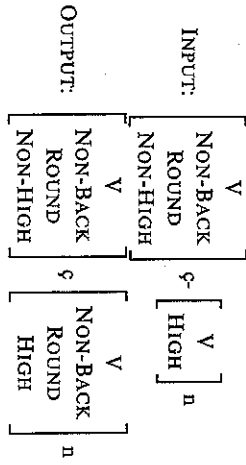
6.23   Output of the phonology for öçler

$$\text{OUTPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}\ \text{çl}\ \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix}_r$$

A mapping from the combination of stored forms to the output generated by the phonology is called a derivation. For the genitive of the same root we have the following derivation, again with missing values copied from the left:
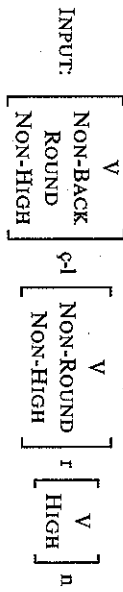
6.24   Derivation of öçün

$$\text{INPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}\ \text{ç-}\ \begin{bmatrix} V \\ \text{HIGH} \end{bmatrix}_n$$

$$\text{OUTPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}\ \text{ç}\ \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{HIGH} \end{bmatrix}_n$$

Now, let's see what happens when we add both suffixes to the same root. Now we have two vowels in a row that are missing values for some features.

6.25   Input representation of öçlerin

$$\text{INPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}\ \text{ç-l}\ \begin{bmatrix} V \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix}_r\ \begin{bmatrix} V \\ \text{HIGH} \end{bmatrix}_n$$

We assume that a vowel that is missing values for any feature pair like ROUND/NON-ROUND looks to its left and copies the first value it finds. Both suffixes have to look all the way to the root vowel to get a value for NON-BACK/BACK. This is the only feature missing for the plural marker. The genitive marker, however, also needs a value for ROUND/NON-ROUND. In the genitive singular, the first such value to the left was on the root, but, in this case, the plural marker has a relevant feature, and the genitive suffix copies it. Thus, we end up with the following:

6.26

Output representation of *öğlerin* with features for the vowels

$$\text{OUTPUT:}\quad \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{ROUND} \\ \text{NON-HIGH} \end{bmatrix}_{çl} \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{NON-ROUND} \\ \text{NON-HIGH} \end{bmatrix}_{r} \begin{bmatrix} V \\ \text{NON-BACK} \\ \text{NON-ROUND} \\ \text{HIGH} \end{bmatrix}_{n}$$

We now see that we can derive the genitive alternations in the singular and plural from the same abstract stored form. There are fewer surface forms of the genitive suffix in the plural because the genitive suffix always gets its ROUND/NON-ROUND value from the plural suffix. The latter is always NON-ROUND since it is underlyingly specified as such. We also see that the form of the suffixes is determined phonologically—the computations depend on the preceding vowel, not on what root the suffix is attached to. This is clear, because different forms of the genitive surface with a given root, depending on what intervenes. The root *öç-* takes a genitive suffix *-ün* if the suffix follows the root directly, but it takes the form *-in* if a NON-ROUND vowel like that of the plural comes between the root and the genitive.

Now, we have just illustrated the computational nature of vowel harmony in Turkish. Getting back to our story, it turns out that the vowels of Turkish roots tend to be harmonic as well, especially with regard to the features NON-BACK/BACK. Here are some examples:

6.27

Harmonic Turkish roots

BACK	boru	"pipe"
	arı	"bee"
	oda	"room"
NON-BACK	inek	"cow"
	dere	"river"
	güzel	"beautiful"

There are some exceptions, especially in recent borrowings like *pilot* "pilot," but for the most part all the vowels in a root will be either NON-BACK or BACK. This is how Charles knew that *Ozel* was a very unlikely name, and that *Özel* would conform to the general pattern of the language.[20]

[20] Paul did manage to visit our venerable doctor, who inadvertently squirted some anesthetic into his eye, and then diagnosed him as having a case of trenchmouth that eventually cleared up on its own. This story is not meant to denigrate the Turkish medical profession—after all, Dr. Özel worked at the *American* hospital in Istanbul.

So, like our analyses of syntactic structure and the allophones of /t/, we once again end up positing very abstract symbols—this time the partially specified vowels of Turkish—to account for complex sets of linguistic data. We turn in the next section to yet one more analysis, from the domain of morphology, that demonstrates just how abstract the equivalence classes are over which grammatical computations apply.

6.5 Words are not derived from words

Recall the discussion of Warlpiri and Samoan from Chapter 1. In those examples, we assumed that the plurals were built from the singulars. Similarly, we saw that the Turkish nominative plural, genitive singular, and genitive plural seem to be formed from the nominative singular form by adding suffixes and providing them with the appropriate values for missing features. So far, it always appeared to be the case that complex words are built out of simple words. However, in some languages, even for simple categories like singular and plural, it is necessary to recognize a complication—it may be the case that both singular and plural need to be expressed as a function of some unit that, by itself, is neither singular or plural. In fact, this unit cannot ever appear as a word on its own. An example will help to clarify.

Here are some singular/plural pairs from the language Swahili, spoken in Tanzania and other countries in eastern Africa (the letter š is pronounced like *sh*). Swahili is the native language of about one million people, but is used as a second language by about 30 million.

6.28

Swahili singular/plural pairs

singular	plural	meaning
mtoto	watoto	"child/children"
mtu	watu	"person/people"
mpiši	wapiši	"cook/cooks"
mgeni	wageni	"stranger/strangers"

Note that each singular/plural pair shares something—for "child/children" it is *toto*; for "person/people" it is *tu*, and so on.

If we want to revive our mathematical analogies from Chapter 1, where we expressed the plural in Warlpiri as a function of the singular, then we can treat the Swahili singular and plural forms as being generated by two different functions of the same independent variable. If we refer to the