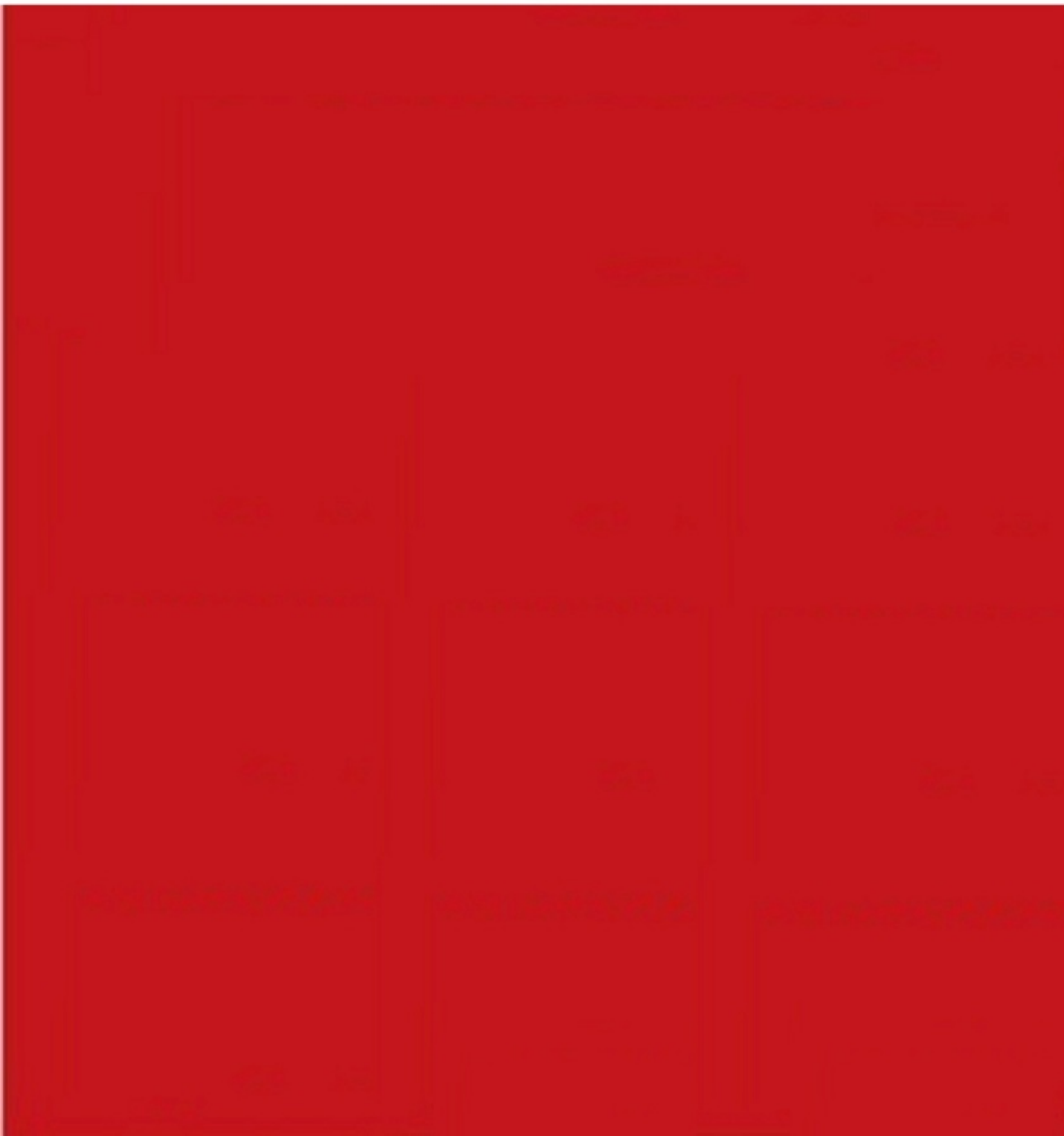
A GRAMMAR OF TURKISH

(A Linguistic Grammar)

Daniel W. Hieber



A GRAMMAR OF TURKISH

by

Daniel W. Hieber

The College of William & Mary
Prof. Reed
Descriptive Linguistics

Fall 2007

Table of Contents

§1. Language Profile

Turkish is a Turkic language (in the Altaic branch) spoken by some 50 million people, the majority of which live in Turkey, where it is the official national language. It is related to some Chinese and Russian languages. The language is divided into eastern and western dialects (the western portion being composed of a single dialect, Danubian). There are 36 total languages existent in Turkey, and of these two are extinct. This total also includes two different versions of Arabic, and Farsi. Turkey itself is divided into Europe and Asia, with 24 and 12 languages respectively.[1] Many of the other Turkic languages in the area form 'chains' of dialects, with nearby languages being extremely similar (and sometimes even mutually intelligible), and more distant ones being decidedly distinct. This continuum of language makes it difficult to define clear language boundaries (Kornfilt, 1987, pp. 619–620). Turkey also includes a number of refugees from Central Asia, numbering in the hundreds of thousands.

The Altaic branch of languages of which Turkish is a part are distributed throughout East, North, Central and West Asia and Eastern Europe. The validity of the language group is sometimes debated, and sometimes Korean or Japonic languages are included in this family. Its location and the earlier expansion of the Ottoman Empire has exposed the language to a number of Persian and Arab influences; a sizeable portion of the vocabulary comes from Arabic, including a few primarily closed-class morphemes like complementizers and conjunctions.

The informant has also provided us with useful knowledge on the background of Turkish. While she was unsure about the status of other languages in Turkey, she knows that there are numbers of Kurds in the country (implying a Kurdish language), and talks about there being other 'provincial' dialects which are nonstandard. She notes that different social classes use the language differently: the upper classes use language more formally (as would be expected), while the lower classes are informal in their language use, and slur words. The informant compared this to the use of the word *y'all* in English, and went on to say that it was "hard to know where one word ends and one word begins." (NihanOnTurkish) She also notices that different regions and dialects depend on the inhabitants in the area, and has noticed these regional differences even within her own family.

---

[1] The preceding information was obtained from the Ethnologue, available at http://www.ethnologue.com.

Turkish underwent significant changes in 1923 as the language was reformed from an Arabic to a Latin alphabet, and generally 'modernized'. However, many pre-reform, archaic words can still be found in the literature. Perhaps because of these reforms, Turkish today is still largely phonetic, and the informant notices almost no code-switching during her time in the country.

Finally, English has a strong presence in Turkey, because it is taught in progressively increasing amounts throughout school, beginning in pre-school, so that by the time one reaches university level, nearly all instruction is in English.

§2. Informant Background

2.1 Personal History

Nihan Kaya, a native first language speaker of Turkish, was born in Ankara, Turkey on May 2, 1988. She is currently nineteen years old, and a student at the College of William & Mary in Virginia. Nihan lived in Ankara until age four, when her father, a software engineer, was offered a job in the state of Pennsylvania. Nihan attended pre-school in the United States. Shortly thereafter she and her family returned to Ankara, where they remained until she was ten years old. Her family returned to the States, living in Connecticut until she had finished sixth grade (at the age of twelve or thirteen). In recent years, she has not had the opportunity to visit Turkey often: her last visit was two years ago, and the previous one another two years before that.

2.2 Family Background

Nihan's immediate family consists of her father, mother, and one sister, 3 years her younger. Her father grew up in southeast Turkey, and they still make the occasional visit there. A nonstandard dialect of Turkish is spoken there, one which is seen as socially backward. Her father reverts to this dialect when he becomes particularly emotional. Nihan says she would often consult her father for English words when she began learning the language. Her father speaks both English and Turkish fluently. Nihan's mother speaks only 'halting' English, and fluent Turkish, and is also from Turkey. Her sister began learning English in pre-school and has had continual exposure to English, and consequently now has some issues speaking Turkish.

Turkish is the primary language spoken in the household, although the 'amount' of Turkish in their family is lessening. Nihan's parents, however, feel as though Nihan and her sister should retain both their language and their culture. Nihan herself agrees, but has reservations concerning the practicality of passing the language on to kids of her own. Her father has little opportunity to speak Turkish outside the household, and no

Turkish social network; her mother, on the other hand, has many friends with whom she can speak the language.

With the beginning of this class, however, Nihan's mother has been excited about her beginning to delve into the language again, and Nihan herself has begun reading books in Turkish again and greatly appreciated the linguistic reawakening, so to speak, and this has made her much more conscious of her own language, and more desirous to preserve it.

2.3 Linguistic History

As already mentioned, Nihan began learning English in pre-school, but she entered standard ESL classes upon returning to first grade in Turkey, in what she describes as a 'normal situation' for the educational system (she was also learning how to read and write in Turkish that same year). This influenced her Turkish from an early stage. Specifically, she cites having had difficulties articulating the vowels of Turkish shortly after her return from the States. Fortunately, as this was still during the critical period, she quickly regained competence in Turkish.

In sixth grade, Nihan began learning Spanish, stopped for a year, then took German for a year, and finally returned to Spanish upon reaching high school, completing through Advanced Placement Spanish (i.e. five total years of Spanish instruction). Nihan also claims to have some comprehension ability in Japanese, due to its syntactic similarity to Turkish, and acquired primarily through a friend who knew the language well. Finally, Nihan speaks fluent English. While there is some clear phonological and semantic influence from English, its effect on her communication is minimal.

2.4 Linguistic Competence

Nihan has very few opportunities to exercise her Turkish on campus, and she has mentioned no other occasions in which she uses the language. It can be assumed then that her use of Turkish is limited to the home. Despite this, she dreams partially in Turkish, and even recounts instances of code-switching in her dreams. At school, she thinks primarily in English, due to the simple fact that she often doesn't know the Turkish equivalent for certain words, although (perhaps obviously) she thinks more in Turkish when around other speakers.

Nihan's self-assessment of her reading comprehension puts her at a fifth-grade level, and she states that she reads the language very slowly. It is also apparent, based on her description of her abilities, that she lacks the higher formality level/register in the language. This she attributes partially to the linguistic reform of Turkish in the 1920's, as many of the archaic (pre-reform) words are still used in the literature, whereas her command of the language is post-reform. Still, she will often read books

in Turkish which she has already read in English, thus abetting her comprehension ability.

There is also talk of what she and her family call 'Turklish', or code-switching into English due primarily to gaps in their Turkish lexicon. Also under her definition of Turklish are instances where the Turkish vowels become marred and acquire qualities of English ones.

Finally, Nihan tells us that she sometimes speaks Turkish too formally at home, something that particularly her family notices. This is actually the opposite of what one might expect, seeing as her Turkish is limited mostly to the (informal) context of the home. At the same time, lack of exposure to Turkish-speaking peers could be a limitation to her acquisition of more casual registers.

2.5 Linguistic Attitudes

Nihan identifies very positively with her linguistic heritage, clearly taking pride in it. In her words, it gives her a sense of 'uniqueness and positivity', as well as a sense of belonging in Turkey. Other practical benefits she sees are extended travel opportunities, and the ability to be discrete via code-switching when necessary.

During elicitation sessions, Nihan has a tendency to assume that there is a 'correct' form of each of the words, based loyally on the orthography of Turkish, which is highly phonetic. This gives our transcriptions a somewhat prescriptivist bent, and perhaps overlooks certain phonetic details (e.g. heavy aspiration on certain consonants, and the varying nature of the 'soft g' which we transcribe as /ʜ/). Additionally, Nihan makes the distinction between 'proper' speech (i.e. enunciated speech) and more improper and natural speech (i.e. conversational speech). Fortunately, she typically gives us both versions whenever this is the case. Not all is hard and fast, however, as she often gives us several variants of the same sentence, telling us which seems the most natural but making little judgment otherwise.

Overall, Nihan's personal and linguistic background, attitudes towards language, and pleasant, energetic personality make for a great informant.

§3. Phonemic Analysis

The following is a list of words elicited from our informant, which serve as the basis for our phonemic analysis of Turkish. Words in which the phonetic and phonemic transcriptions are interestingly different are noted. These are instances in which my or other students' transcriptions deviate from the official data. Otherwise, all phonemic transcriptions are those adopted and standardized by the class.

3.1 Data

Transcription	Translation
[abɪ] /abi/	older brother
[a:tʃ] /aʜatʃ/	tree
aʜɨr	heavy
ahɨr	filth, pigstye
[a:iz] /aʜɨz/	mouth
ai	moon
[aˑjak] /ajak/	foot
akʃam	night
anna	mother (informal)
annɛ	mother (formal)
annɛannɛ	mat. grandmother
atʃ	open
[ʌtɛʃ] /atɛʃ/	fire
[bʌba] /baba/	father
[bʌdʒak] /badʒak/	leg
baʃ+parmak	thumb
[bʌlɨk] /balɨk/	fish
-bɛ	and (enclitic)
[bæn] /bɛn/	I, myself
Biz	us
buz*	ice
boʜaz	throat
bozuk	rotten
[bulutʰ]/bulut/	cloud
[bu:run] /burun/	nose
[bʲykannɛ] /bykannɛ/	pat. grandmother
[bʲykbʌba] /bykbaba/	pat. grandfather
/dɛdɛ/	mat. grandfather
dʒan	life
ɛʃ	mate, spouse
ɛk	bury
ɛl	hand
ɛm	suck
ɛn	with
ɛr	soldier
[ɛtʰ]/ ɛt/	meat
ʃu	this
[ev] /ɛv/	house
[gɛtʃɛrɪm] /gɛtʃɛrim/	I will pass
[gʲøz] /gøz/	eye
[kyn] /gyn/	day
gynɛʃ	sun
[hʌva] /hava/	sky
[ɪnsan] /insan/	human
jaʜlɨ	greasy, fatty
[ja:mur] /jaʜmur/	rain
japrak	leafy plant
[jɛ]	eat
jɛʃil	green
jɨlan	snake
jɨldɨrɨm	lightning (noun)
jɨldɨz	star
joʜurt	yoghurt
jumurta	egg
kaʃ	eyebrow
kafa	head
kaj	slide
[kaˑja] /kaja/	rock
kal	stay (imp.) stale, adj.
kan	blood
kap	cup
[karza] or [kar]	snow
kardɛʃ	younger brother
karpuz	watermelon
kat	floor of building
katʃ	run away (imp.)
[kɛmɪkʰ]/ kɛmik/	bone
[cɪn] /kin/	spite

kon*	spike
[kʲirlɛ] /kirlɛ/	dirty (formal)
kɨrmɨzɨ	red
kɨsa	short
kɨz(ɨ)	girl (acc.)
kod	code
kol	arm
kom	land (v.)
[kop̚] /kop/	snap
[kʰøp̚ek] /køpɛk/	dog
kor	fire (big fire)
kot	denim
kotʃ	ram (animal)
koz	trump
kuʃ	bird
kul	person, soul
kum	sand
kyl	ashes
lahana	cabbage
[mavɪ] /mavi/	blue
[mutlo] /mutlu/	happy
nɛ	what? (informal)
oha	*&!#, expression of surprise, like 'damn!'
[okianus] /okjanus/	ocean
[oːtʰ]*	weeds/grass
parmak	finger
[pɪs] /pis/	dirty (informal)
rus*	a Russian
[ɹysgʲar] /ryzgar/	wind
saʜlɨk	health
saʜol	thank you
sarɨ	yellow
satʃ	hair (on head)
saz	a Turkish instrument
[sɪz] /siz/	you (singular, formal)
[soːuk] /soʜuk/	cold (adj)
soludʒan	worm
søz	promise
[suː] /su/	water
[ʃuː] /ʃu/*	this
sus*	be quiet! (rude)
[sytʰ]/syt/	milk
taʃ	stone
tʃim/tʃim/	lawn grass
[tʃɪtʃɛk]	flower
[tɛçlike] /tɛlikɛ/	danger
tʃodʒuk	child
tʃul	clothes, rags
toprak	earth
[tuz̚]	salt
utʃurum	cliff
uzun	tall
vadi	valley
van	a city in Turkey
vɛ	and

*Word not listed in the class data set.

3.2 Transcription Discrepancies

I do not feel the need to offer alternative transcriptions for any of the words in the class data set. In all cases but one, I agree that the class data is the correct phonemic transcription of the word. Where my transcriptions differ, however, are in the phonetic transcriptions. I simply note a somewhat narrower transcription in a number of cases.

So while I am in complete agreement with the phonemic transcriptions for our data, my only points of contention are the phonetic realizations of those phonemes.

Some common examples of places where I have indicated a phonetic realization that is different from the phonemic sound are: the use of [ʌ] for a stressed /a/; the phoneme /i/ being realized as [ɪ]; noting obvious aspiration; and inserting the palatal [j] where it may have been heard during elicitation. But again, I don't think that any of these are phonemically salient.

The one word where my transcription differs from the class set is with /ɛr/ 'soldier'. I heard this clearly as [ar], yet even the informant insisted on the vowel quality of the epsilon. So again, I concede that the phoneme in use here is /ɛ/, but simply being realized as [a].

While I have decided to adhere to the standard set forth by the class for our data, I would offer one minor notational suggestion: as will be shown later on, both [ɪ] and [i] are conflated into (what we have decided to call) /i/. However, in the large majority of cases, /i/ is realized phonetically as /ɪ/ instead of /i/. This leads me to think that /ɪ/ serves as a better representative symbol of the phoneme than /i/. However, as will also be shown later on, Turkish has no tense/lax distinction in its vowels, and the phonemes we have chosen to use are all lax. So the use of /ɪ/ as a phoneme would force us to reconsider – or at least be cautious in our approach to – our judgement about tense v. lax vowels. Note that if all the front vowels were tense, and back vowels lax (as there is some evidence suggesting), then the contrast between tense and lax wouldn't be psychologically salient, and could be subsumed under +/- frontness instead.

3.3 Transcription Difficulties

Our most difficult transcription problem involved identifying and learning to recognize the presence of a 'mystery sound' which our informant describes as a "soft *g*." Nihan had to help us a great deal in the study of this particular sound. The sound *did* make itself salient to us however, independent of Nihan's insistence on its presence. Her remarks were triggered precisely because we were beginning to notice something, and needed assistance in figuring it out.

The soft *g* stood out to us originally precisely because we couldn't process it as a sound we knew. Mostly, it sounded like lengthening on the vowel. At various points in our early transcriptions, we had transcribed the sound as a glottal stop, vowel lengthening, or an extra vowel. So while Nihan had to tell us explicitly where this sound was, and how we had been mistranscribing it, the fact that the sound surfaced in some form or another in our transcriptions before her intervention shows that it is in

fact a tangible sound in the language, and not simply a psychological representation in the informant's mind.

Other issues for the class included discerning the voicing quality of certain consonants – namely /k/ versus /g/ and /s/ versus /z/. The voicing distinction between these was often very subtle, and difficult for us to discern. This problem was especially exacerbated in instances where voicing/devoicing based on the environment may have been occurring. We as a class also had a tendency to hear /l/'s and /j/'s where they didn't belong, due to the quality of the nearby vowels. Usually, these could be attributed to the position of the vowel formation in the mouth (e.g. all instances of suspect palatal /j/'s seem to be cases where the /j/ is followed immediately by a high vowel). Aspiration also threw us off at times: the aspiration on certain consonants is so strong that, a number of times, we initially transcribed the words with extra consonants that were in fact just voicing (e.g. originally /buluts/ for [bulut^{h}]).

A problem for us as English speakers was the conflation of certain vowels in Turkish, which we distinguish in English. It seems from the data that there is no tense/lax distinction in Turkish, though the distinction is obvious to us. Many students wanted to transcribe different phones as separate phonemes, particularly in the case of [ɪ] and [i]. However, the data better supports the hypothesis that the two are simply different phonetic realizations of the same phoneme /i/.

Generally, our transcription symbols began to standardize as we encountered certain sounds more frequently and realized which ones were phonemically salient. At first, the process was made difficult by the fact that the informant would occasionally pronounce sounds with subtle differences. This only makes sense – the speaker is not going to pronounce the same word identically every time she utters it. Eventually, we were able to get enough of an 'aggregate' of any one sound to realize which variations in pronunciation were simply 'outliers'.

In terms of elicitation techniques, there were no real issues to speak of. Nihan is an excellent informant, and even returns to us with corrections on her own data, or corrections on our missteps. There were a number of techniques that served us particularly well:

- Asking the informant if sound x was the same as sound y in a different word
- Asking the informant if the vowels in a word were all the same, or if some were different
    - In retrospect, this is especially useful considering the vowel harmony in Turkish. Without even realizing it, we were utilizing Turkish's vowel harmony with our informant to determine certain vowel sounds
- Repeating a word back to the informant in two different ways, and asking which sounded closer

- Asking the informant to say two words with similar sounds to compare/contrast those sounds
- Making novel sound combinations in hopes of finding minimal pairs
    - This turned out to be an extremely useful technique, eliciting much more data than I had expected

Even given this, there are some inherent issues with the elicitation process. First is the general issue of translation, rather than spontaneous speech. The conceptual schema of English may be a significant influence on the sentences elicited. Second, the speaker's judgments will vary from day to day or even sentence to sentence. She may feel in a more formal mood one day, and a more casual one the next. Finally, there are always the random bits of miscommunication which plague any one of our sessions.

3.4 Phonemic Inventory

Figure 1 shows an inventory of the vowels in Turkish, based on phonemic contrasts found using minimal pairs in the language. /ə/ or /ʌ/ also occur, but there is no basis for considering them as distinct phonemes. In all these cases, the conferment of the status of 'phoneme' on these sounds is directly motivated by strong contrasts using minimal pairs (the data of which will be discussed in the next section).

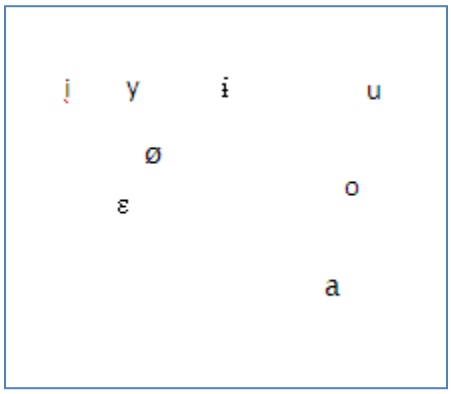


Figure 1

	+F		−F	
+H	i	y	ɨ	u
−H	ɛ	ø	a	o
	−R	+R	−R	+R

Table 1

One of the first things to note from this vowel chart is that there is no tense/lax distinction in the language. Thus the class has commonly heard sounds similar to both [i] and [ɪ], for example, which the informant hears only as /i/. The same holds true for [e] and [ɛ]. Rounding also appears in more places than English, with rounded vowels both high and front. Generally then, there are two basic vowel contrasts for speakers of Turkish: front v. back, and rounded v. unrounded, in place of a tense/lax distinction.

Table 2 shows the consonant inventory for Turkish, again based on phonemic contrasts garnered from minimal pairs. Each sound was compared with any sounds that were phonetically similar, as finding minimal pairs for entirely unrelated sounds tells us little of use.

	Bilabial	Labio-dental	Alveolar	Alveo-palatal	Palatal	Velar	Glottal
Nasal	m		n				
Plosive	p b		t d			k g	
Fricative		f v	s z	ʃ			
Affricate				tʃ dʒ			
Approximant			r		j		h ʜ(ɦ)
Lateral Approx.			l				

Table 2

I have decided to analyze our mystery sound – which we chose to transcribe as /ʜ/ – as a glottal approximant, for several reasons. Two things seem obviously clear about the sound: first, it must be an approximant, because the class heard no blockage of airflow during elicitation. This explains why none of us heard this sound as anything resembling a consonant at first. Secondly, the sound is typically voiced, because it falls between two vowels in such a way as to blend with them almost entirely – this is why most of the class hears this sound as nothing more than lengthening. The two choices that seem to fit this sound best should then be the velar /ɯ/ or the glottal /ɦ/. And because the glottal creates a voiced pair with /h/, fulfilling the symmetry that seems present in most of the rest of the consonant inventory, the glottal seems a better choice than the velar approximant. Still, this sound clearly has several different realizations depending on its environment, so it is difficult at best to assign any one phonetic value to it.

In contrast to Figure 1, Table 1 is a phonemic chart of the vowels in Turkish. The pivotal difference here is that, while Figure 1 simply shows the phonetic value of the phonemes in Turkish, Table 1 organizes these phonemes according to the features which are salient to a speaker of the language. We know from vowel harmony that the relevant features are R(ounding), H(eight), and F(rontness), and that vowels harmonize in accordance with these features. There are slight differences between the two charts as well. In the phonemic chart, the /a/ phoneme is realized psychologically by the speaker as -R, even though it is phonetically +R. Also, /ɨ/ is more naturally /ɯ/, because it fits the symmetry of the system best. Despite this, we continue to use /ɨ/ to represent the sound, whatever it may be phonetically.

3.5 Minimal Pairs

Towards the end of our elicitation of single words, we discovered the presence of vowel harmony in Turkish, based primarily on the features of roundness or front-/backness, or both. Unfortunately, this greatly complicates our search for contrastive vowel sounds, for the simple reason that, if one vowel in a word changes, the vowels around it will have to change as well in order to fit the pattern of vowel harmony. So the odds of finding a minimal pair in a multisyllabic word are next to impossible.

With that said, we are still perfectly capable of establishing the phonemic reality of all our vowels based on strong, monosyllabic minimal pairs. Table 3 shows these contrasts.[2] The table is extremely conservative, including only strong instances of contrast. However, there was plenty of more weakly contrasting data that further supported these distinctions. For example, a strong contrast exists between /a/ and /u/, based on several minimal pairs (e.g. /kaʃ/ versus /kuʃ/). Yet there are other instances (such as /ʤan/ versus /ʤuk/) which contrast similar-enough environments to lend good support to the reality of a phonemic contrast.

	a	i	ɨ	ɛ	u	ø	y
a							
i	kan / kin						
ɨ	kal / kɨl	–					
ɛ	anna / annɛ	–	–				
u	kaʃ / kuʃ kal / kul	siz / sus biz / buz	kɨl / kul	kɛmik / kum			
ø	saz / søz kap / køpɛk	–	–	–	sus / søz		
y	kal / kyl	kin / gyn	kɨl / kyl	–	kul / kyl rus / rysgar	–	
o	kan / kon kal / kol kap / kop kat / kot katʃ / kotʃ	kin / kon	kɨl / kol	ɛt / ot kɛmik / kom	kul / kol kum / kom	gøz / koz køpɛk / kop	gyn / kon kyl / kol

Table 3

Interestingly, the vowel harmony itself can be a useful tool for establishing contrast. Because changing one vowel affects the harmony of the word as a whole, we should expect to see contrasting harmonies, after a fashion. This is in fact the case: the informant clearly contrasts /varmuʃ/ with /vɛrmiʃ/, for example, which is strong evidence supporting all four vowels as independent phonemes.[3] Basically, this means that pairs of vowels are more likely to contrast than individual ones – at least in multisyllabic words.

---

[2] Some cells are left blank to avoid redundancy.

[3] We know that Nihan contrasts these – specifically the /a/ and /ɛ/ – because when one student mispronounced /vɛrmiʃ/ with an /a/ as the first vowel, the student was corrected by Nihan saying, 'No, that would be /varmuʃ/, which is a different word.' Despite the fact that the student pronounced only one of the vowels wrong, Nihan instantly heard *both* vowels as being fully contrastive with /varmuʃ/.

	p	b	f	t	d	r	l	n	ʧ	s	m	v	ʤ	g	h	ʜ	j	ʃ	z	k
p	–																			
b	biz / pis	–																		
f	kafa / kap		–																	
t				–																
d				kod / kot	–															
r				ɛr / ɛt kar / kat rus / tuz		–														
l				ɛl / ɛt kal / kat kol / kot		ɛl / ɛr bulun / burun kal / kar kol / kor	–													
n				ɛn / ɛt kon / kot				–												
ʧ				katʃ / kat kotʃ / kot					–											
s				–					–	–										
m		–						ɛm / ɛn kom / kon			–									
v		mavi / abi vɛ / bɛ hava / baba van / bɛn	hava / kafa									–								
ʤ					–				ʧoʤuk / ʧiʧɛk				–							
g														–						
h														–	–					
ʜ														–	aʜir / ahɨr	–				
j					jɛdi / dɛdi												–			
ʃ									kaʃ / katʃ ʃu / ʧul	ʃu / su								–		
z									saz / satʃ	kɨza / kɨsɨ			–						–	
k														koz / gøz kum / gyn	kafa / hava					–

Table 4

Consonants were generally easier to find contrasts with, though of course monosyllabic minimal pairs were again the most useful, and easiest. Again, this data is supplemented by many more cases of contrasting environments, which are a weaker contrast than minimal pairs, where the contrast is so salient that it brings about a change in meaning in the word. Table 4 shows contrasts between consonants with similar articulation (consonants with little or no phonetic similarity were not sought after, and those cells are left blank; cells with a dash indicate that no minimal pairs were found).

There were some gaps in the data where I would have preferred to see a strong contrast instead. Despite this, I feel as though all the phonemes listed in the inventory in Table 1 are adequately motivated.

To start with, there were no contrasts between /d/ and /ʤ/, not even in similar environments. Only a few weak contrasts were established: a V_V contrast (va<u>d</u>i / solu<u>ʤ</u>an), a #_V contrast (<u>d</u>ɛdɛ / <u>ʤ</u>an), but not a V_# contrast. However, we can still motivate the contrast between the two sounds by symmetry to the contrast between /t/ and /ʧ/ – we should expect the voiced counterpart. Likewise, the data for both /d/ and /ʤ/ is very limited, so I think more elicitation would highlight this contrast in the future.

Next, /j/ had only one minimal pair (j<u></u>ɛdi / <u>d</u>ɛdi), and /j/ and /d/ are phonetically similar, so this helpful in establishing /j/ as a phoneme. The informant also seems extremely confident of /j/'s status as a phoneme, overtly referring to the sound any number of times during our elicitation. Plus, the specific minimal pair just cited happens to be two variations on the same word, the difference being one of conjugation on the verb. This means that the /j/ here doesn't just contrast with /d/ – it is in fact has meaning as a grammatical morpheme. So despite a lack of contrasts, I think there is little reason to discredit /j/ as a phoneme. One potential problem, however, is distinguishing its use as a consonant from its use as a vowel – our transcription has not been consistent on this point so far. It may be that there are phonological guidelines – perhaps about the syllable structure – which determine /j/'s status as either a consonant or vowel.

Another very bothersome gap is the lack of contrast between /g/ and either /h/ or /ʜ/, especially since they are so close phonetically – the informant even calls our /ʜ/ a 'soft *g*'. My instincts, and that of the informant, would tell us that all three of these are separate sounds, however at the moment the data may not warrant it. This is simply because there is very little data with these sounds. There are, however, contrasts between /h/ and /ʜ/, and between /g/ and /k/, which means all three of the sounds in question should be able to establish their phonemic status independently of each

other. So again, I think future elicitation will further support a phonemic difference between these sounds.

3.6 Syllable Structure

Stress always falls on the first syllable in Turkish. One extremely common word structure is CVCVC, and in all these cases, my judgement is that the middle C belongs to the second syllable, and not the first, and this may somehow be related to the initial stress on the word. The exception to this is in words with two medial consonants, in which case the first consonant often belongs to the first syllable, and the other to the second. Of course, other times the split is CV/CCVC (e.g. *toprak*). That structure – CVCCVC – is also very common in our data. Finally, the basic pattern of CVC is practically ubiquitous. Between these three structures, I estimate that about 80-90% of all our words are accounted for.

Compare the four paradigms, divided by syllables:

CVC
CV/CVC
CVC/CVC
CV/CCVC

From just this data alone, one would strongly suspect a CVC syllable structure as being the most basic to Turkish. In fact, one can posit a single paradigm that accounts for almost all our data: (C)V(C)/CV(C). This is not a very powerful observation, however. It is merely inferring a pattern retroactively from the data. Luckily, we are able to make the move to a slightly more explanatory – and thus more powerful – observation.

Starting with the obvious inference that CVC seems to be the basic syllable structure, it's worth noting that the first C of the second syllable in the above paradigm is interesting because it is the only one that isn't optional. This suggests that speakers of Turkish have a strong tendency to avoid vowel clusters. The informant was certainly right in her claim that vowels never stand next to each other in Turkish – they are usually separated by a consonant or approximant, however subtle (as in the case of /ʜ/). This can be a very useful thing to keep in mind during our transcriptions, especially when listening for /ʜ/, or when deciding between cases like /boɪ/ versus /boj/ (which inadvertently helps motivate the phoneme /j/ as well).

Other evidence for a basic CVC structure is that the only multisyllabic word beginning with a vowel in our data is the loanwoard /okjanus/ 'ocean', and that even the monosyllables are overwhelmingly CVC.

The same tendency to avoid vowel clusters also helps explain the presence of consonants at the beginning and end of most Turkish words. Because several of the morphemes we have encountered begin with (or consist entirely of) a harmonized vowel, consonants at the beginning or end of words helps prevent vowels from

meeting at the word boundary. Occasionally, the job of separating vowels is accomplished by consonant epenthesis instead (usually a /j/), but CVC the syllable structure certainly abets the avoidanc of vowel clusters. Likewise, even the slightly more complex morphology we have seen follows the CVC pattern (specifically -muʃ). Generally, grammatical morphology and enclitics will be as reduced as possible, and most often consist of just a single syllable. So the fact that the grammatical marker -muʃ supports a CVC structure.

Finally, our mystery sound /ʜ/ seems particularly well-adapted to separating vowels, as it is frequently intervocalic, and seems to require being preceded by a vowel. So it looks as though Turkish needs some sound – however subtle, as in the case of /ʜ/ – to separate its vowels, and this is the driving force behind its CVC syllable structure. This in turn provides us with an explanatory theory, rather than just an observational one concerning Turkish syllable structure.

3.7 Phonological Rules

Turkish commonly alternates between /ʜ/ and /k/, and other voiced/voiceless stop pairs, depending on their position in the word. In all word-final positions, stops are voiceless, suggesting word-final devoicing. In intervocalic or word-medial positions, however, the phoneme /k/ is realized as /ʜ/ (kopɛk v. kopɛʜi), and likewise /ʧ/ as /ʤ/ (anaʧ v. anaʤɨn), and /p/ as /b/ (kitap v. kitabɨ). In one interesting case, the citation form of a word (/daʜ/ 'mountain') ended in the voiced /ʜ/, whereas typically the citation form would be /dak/, suggesting a process of word-final devoicing, rather than intervocalic voicing. There are also numerous instances of voiceless stops between vowels (e.g. /oku/).

Perhaps related to this is the fact that many voiceless stops have heavy aspiration, especially in the devoiced positions. So a quick glance at the data at the beginning of this section will show that almost all word-final stops have heavy voicing, and even /r/'s have some type of frication in a way that seems likewise related. Other stops are swallowed or unreleased, especially intervocalically.

The vowels /ɛ/ and /i/ also undergo significant variation depending on their environment. For /ɛ/, /ɛ/ → /æ/ | _/c̃/, and /ɛ/ → /e/ perhaps before an /r/. In the plural morpheme, for example, /lɛr/ sounds extremely close to [lar], making it hard to distinguish from /lar/. It also explains my earlier comment regarding hearing /ɛr/ 'soldier' as [ar]. Interestingly, /i/ is realized in a manner somewhere between [ɛ] and [i] when word-final. In the majority of cases, /i/ is realized as [ɪ].

There are two 'rules' which relate to the way in which words and morphemes undergo vowel harmony. The first, which I call Harmony Rule #1, harmonizes a vowel with the nearest preceding vowel. The harmonized vowel will be -R, -H, and agree with

the preceding vowel in frontness. As an example, the plural morpheme {-lVr} takes this harmony, so that when preceded by a /i/, /y/, /ɛ/, or /ø/ it becomes /lɛr/, and when preceded by /ɨ/, /u/, /a/, or /o/ it becomes /lar/.

The second harmony rule, taken by certain morphemes like the possessives and seen in the examples given in the first paragraph of this section, sits in opposition to the first in that the harmonized vowel is +H, and agrees with the quality of the preceding vowel in frontness and rounding. Under this rule, /ɛ/ and /i/ give rise to /i/, /ø/ and /y/ to /y/, and so forth.

These harmony rules hold internally within the stem of most lexemes, and I assume that any which do not adhere to these harmony rules are loanwords historically. Affixing morphemes (such as case marking) tend to code for one or the other type of harmony, although in a few instances there is a stable vowel in part of the morpheme. Vowel harmony also only occurs over the breadth of a single word. This simple fact is useful in determining word boundaries.

3.8 Language Universals

Turkish seems relatively uninteresting as far as language universals go. Still, it is worth comparing Turkish to the generally accepted linguistics universals.

First, consider the non-absolute universals: most languages have twice as many consonants as vowels, and Turkish fits this nicely – 20 consonants and 8 vowels. They also have generally between 20 and 37 segments – Turkish has 28. It has six different stops, while most languages have at least one. If a language has voiced stops, it must also have voiceless ones, and the majority of languages have more voiceless than voiced stops. Turkish has both voiced and voiceless stops, however it has exactly three of each, rather than more voiceless ones. Like most languages, the stops are /p/, /t/, and /k/, along with their voiced counterparts.

Languages typically have up to 4 fricatives, though some can have much more. Turkish has 6, two of which are voiced counterparts of the voiceless fricatives, in effect making exactly 4 places of articulation. There are several universal implicatures, listed below, which Turkish also follows:

If /p/, then /k/
If /k/, then /t/
If /g/, then /d/
If /d/, then /b/
If /m/, then /n/

These universal implicatures also directly motivate the reality of every phoneme in the ‘then…’ position, further supporting our data. In fact, all these universals indirectly support our data.

Continuing on, Turkish is like most languages in that it has the nasal /n/, which is most common, as well as /m/. 72% of languages have two liquids - /l/ and then some form of /r/ - and this is the case with Turkish as well. /j/ is the most common approximant in languages, and it is also the only true approximant in our inventory of Turkish.

As for vowels, languages usually have between 5 and 7, with the lower limit set at 3. Turkish has 8, which is only slightly more than average. Every language must have a high front unrounded vowel, usually /i/ or /ɪ/; in Turkish, these two vowels appear to be allophones of the same phoneme, seemingly in relatively free variation (though we have yet to test for complementary distribution). Languages must have a low vowel and a high back vowel, with low vowels generally being more central, all of which also hold true for Turkish. Languages with 5 or more vowels generally also have a mid back rounded vowel, which Turkish does (/o/). The one place where Turkish is unusual is in the presence of rounded phonemes both high and frontally, because it is more typical see unrounded front vowels and rounded back vowels.

§4. Morphological Analysis

The following analysis is based on the data from 300 Turkish sentences elicited from our informant by giving the English and asking for the Turkish equivalent. Below is a list of all the word stems (i.e. root lexemes, lacking any additional morphology) encountered during elicitation, categorized according to lexical class.

4.1 Data

Nouns:

abi	older brother (n.)
adam	man (n.)
ajɨ	bear (n.)
ajʃɛ	Ayşe (n.)
akʃam	night (n.)
alt	bottom (n.)
ananas	pineapple (n.)
annɛ	mother (n.)
armut	pear (n.)
aslan	lion (n.)
aʜatʃ	tree (n.)
balɨk	fish (n.)
baʃ	head (n.)
bɛn	I (pro.)
bitki	plant (n.)
bulaʃɨk	dish (n.)
burak	Burak (n.)
dans	dance (n.)
daʜ	mountain (n.)
dɛrs	lesson (n.)
dilini	tongue, language (n.)
ʤuma	Friday (n.)
ʤumar-tɛsi	Saturday (n.)
duvar	wall (n.)
dyn	yesterday (n.)
ɛlbisɛ	dress (n.)
ɛlma	apple (n.)
ɛv	house (n.)
farɛ	mouse (n.)
fil	elephant (n.)
gɛjik	deer (n.)

gøkhan	Gökhan (n.)
gytʃ	strength (n.)
hɛdi	gift (n.)
ismi	name (n.)
itʃ	inside (n.)
jalan	lie (n.)
jardɨm	help (n.)
jarɨn	tomorrow (n.)
jɛmɛk	food (n.)
jɛr	ground (n.)
kadɨn	woman (n.)
kaja	rock (n.)
kalɛm	pencil (n.)
kaplumbaʜa	turtle (n.)
kardɛʃ	sister (n.)
kɛdi	cat (n.)
kɛmik	bone (n.)
kɨl	hair (n.)
kitap	book (n.)
kɨz	girl (n.)
køpɛk	dog (n.)
kurabijɛ	cookie (n.)
kuʃ	bird (n.)
kutu	box (n.)
majmun	monkey (n.)
mantar	mushroom (n.)
marɛ	cave (n.)
masa	table (n.)
mɛktup	letter of correspondence (n.)
o	he/she (pro.)
øbyrsy	day after tomorrow (n.)?
oda	room (n.)
okul	school (n.)
ola	event (n.)
olma	floor (n.)
onlar	they (pro.)
orman	forest (n.)
oʜlan	boy (n.)
øʜr-ɛndʒi	student (n.)
øʜr-ɛtmɛn	teacher (n.)
øʜyn	dinner (n.)
paljotʃa	clown (n.)
pasta	cake (n.)
pazar	bazaar (n.)
pazar(gyn)	Sunday; bazaar (n.)
pazar-tɛsi	Monday (bazaar-after) (n.)
pɛrʃɛmbɛ	Thursday (n.)
rɛsim	painting/picture (n.)
salɨ	Tuesday (n.)
ʃark	song (n.)
saz	Turkish instrument (n.)
sɛn	you (sing.) (pro.)
ʃɛrap	wine (n.)
siz	you (pl.) (pro.)
sokak	street (n.)
su	water (n.)
tabak	plate (n.)
tʃarʃamba	Wednesday (n.)
tavʃan	rabbit (n.)
top	ball (n.)
tula	brick (n.)
tyrktʃɛ	Turkish language (n.)
uju	sleep (n.)
yst	top (n.)
zaman	time (n.)

Verbs:

al	take (v.)
al?	buy (v.)
as	hang (v. tr.)
at	throw (v.)
bɛslɛ	feed (v.)
bil	know (v.)
burak	leave (v.)
dojur	feed (v.)
dyʃ	fall (v.)
gɛl	come from (v.)
gi	wear (v.)
gir	enter (v.)
git	go (v.)
gør	see (v.)
inan	believe (v.)
ɨsɨr	bite (v.)

ist	want (v.)
itʃ	drink (v.)
jakala	catch (v.)
jap	do (v.)
jaʃa	live (v.)
jaz	write (v.)
jɛ	eat (v.)
jɨka	wash (v.)
jola	send (v.)
kɛs	cut (v.)
koj	put (v.)
kokla	smell (v.)
kønuʃ	speak (v.)
kork	fear (v.)
korkut	scare (v.)
koʃu	run (v.)
kVʃ	wish (v.)
oku	read (v.)
øl	die (v.)
øldyr	kill (v.)
onla	realize (v.)
otur	sit (v.)
piʃir	cook (v.)
sɛv	like (v.)
søjlɛ	sing (v.)
tʃal	steal (v.)
tanɨs	meet (v.)
tɛkmɛlɛ	kick (v.)
tɛmizlɛ	clean (v.)
tɛrkɛt	leave (v.)
tʃɨkar	take out (v.)
tʃiz	draw (v.)
topla	pick (v.)
tʃyry	rot (v.)
umu	hope (v.)
vɛr	give (v.)
yz	upset (v.)
yzyl	regret (v.)

Adjectives:

ajnɨ	same (adj.)
bɛjaz	white (adj.)
bɛrabɛr	together (adj.)
byk	big (adj.)
daha	more (q./adj.)
ɛski	past (adj.)
janlɨʃ	wrong (adj.)
jɛni	recent (adj.)
kahvɛ-rɛngi	brown (adj.)
karanlɨk	dark (adj.)
kɛndi	self (adj.)
kɨllɨ	hairy (adj.)
kytʃyk	small (adj.)
mavi	blue (adj.)
øndʒɛ	before (adj.)
sijah	black (adj.)
tɛk	alone (adj.)

Quantifiers:

bir	one (q.)
daha	more (q./adj.)
dørt	four (q.)
iki	two (q.)
ytʃ	three (q.)

Determiners:

bir	a (det.)
bu	proximate (det.)
o	distal (det.)

Interrogatives:

hangi	which? (interrogative)
kim	who? (interrogative)
nɛ	what? (interrogative)
nɛrɛdɛ	where (interrogative)
nijɛ	why? (interrogative)

4.2 Nouns

Of the data set above, 63 words are what we shall call nouns. The majority are concrete entities ('box', 'mushroom'), but the list includes also abstract nouns ('inside', 'strength'), pronouns ('I', 'they'), and proper names ('Burak', 'Ayşe'). Each subcategory of noun has subtle differences in use or meaning, but they all share common characteristics which justify their placement under the category 'noun'.

To begin with, concrete entities, abstract ones, pronouns, and proper names can all function as similar thematic roles – Agent, Patient, or Recipient primarily. Additionally, they fit the prototypical semantic categories of people, places, or things. Syntactically, all of these are words which can be modified by either determiners or adjectives, and may occupy both subject and object positions in the sentence. The entire set of nouns fit basic test frames for both subject and object position (e.g. _____ *kitap okudu* '_____ read the book', *adam* ____ *ʧaldɨ* 'the man read the _____'), including somewhat more complex locative constructions (e.g. *adam kitap kadɨnɨn* _____*-Vnɛ kojmuʃ* 'the man put the book on the woman's _____'). Lastly, these words share common morphology, being able to take marking for plural, possession, and thematic role/case.

The morphology on Turkish nouns follows the pattern below:

STEM + PLURAL + POSSESSION + CASE MARKING

All Turkish morphology undergoes some form of vowel harmony with the vowels of the stem or proceeding morphology. Each affixing morpheme follows one of two rules for vowel harmony: certain morphemes harmonize based on both rounding and frontness, while others harmonize based on frontness alone. We will see this as we examine each morpheme, starting with plural marking.

4.2.1 Plurality

Plurality is marked by affixing the morpheme {-lVr} to the end of the noun, where the harmonized vowel is -H(igh) and -R(ounding), and whose frontness is given by the quality of the nearest preceding vowel (this morphology is also used with verbs, as we will see later). Put simply, if the previous vowel is +F(ront), the vowel in {-lVr} will be an /ɛ/; if it is -F, the vowel will be /a/. From here on, this will be referred to as Harmony Rule #1. Nouns are necessarily *un*marked for plurality when modified by a quantifier; but in all other cases a plural noun will be marked.

4.2.2 Possession

Possession in Turkish is marked on both the thing possessing, i.e. the possessor (POSSR.) and the thing possessed (POSSD.). The POSSR is marked by affixing the morpheme {-nVn-}, where the harmonized vowel is +H, and whose frontness and

roundness is given by the quality of the nearest preceding vowel. This I refer to as Harmony Rule #2. The POSSR morpheme has two allomorphs: /Vn/ and /nVn/. Which allomorph is utilized depends on the preceding phoneme:

$$\{\text{-nVn-}\} \rightarrow \begin{pmatrix} /\text{nVn}/ \\ /\text{Vn}/ \end{pmatrix} \Big| \begin{pmatrix} /\text{C\_}/ \\ /\text{V\_}/ \end{pmatrix}$$

The POSSD marker is primarily just {-V(n)}, though it varies:

(73) aiʃɛ-nin kɛdi-si
Ayşe-POSSR cat-POSSD
'Ayşe's cat'

(74) gøkhan-ɨn kɛdi-sin-ɛ
Gökhan-POSSR cat-POSSD-POST
'to Gökhan's cat'

(69) kɛdi-nin mantar-ɨn-ɨ
cat-POSSR mushroom-POSSD-ACC
'the cat's mushroom' (ACC.)

When the possessor is the third person, the POSSD marker agrees in person, taking the phoneme /s/ to show agreement. In the POSSD morpheme, one allomorph is used word-finally, the other word-medially:

{-(s)Vn-} → /(s)V/ | /_#/

Possession is very basic to Turkish constructions. In addition to simple phrases showing physical possession of an object, possessive constructions are used to show location, and can even affix to interrogatives:

(184) kim-in kɛdi-sin-i ʧal-dɨ
WH-POSSR cat-POSSD-ACC steal-PAST
'whose cat did he steal?'

(59) kitap masa-nɨn yst-yn-dɛ
book table-POSSR top-POSSD-AT
'the book is on the table'

The possessor of a sentence is always implied, so that a sentence can have only a possessed noun:

(90) o kɛdisinɛ jɛmɛk vɛrmiʃ
he cat-POSSD-TOfood give-PAST
'he$_i$ gave food to his$_i$ cat'

4.2.3 Object Marking

I hesitate to analyze the Turkish system of marking its direct object/Patient as a case system, simply because the only case we actually have strong evidence for is the accusative, and potentially a dative. This object marking also appears to be optional in certain contexts, whereas case marking is usually obligatory. It also takes no agreement (i.e. it has no case spreading), at least not on adjectives. Yet motivating the accusative case is fairly easy: to say that Turkish just marks its patient (and *not* an accusative case) would be to overlook several facts. The morpheme (whatever it is meant to represent) appears on more than just the Patient; it can mark the Theme, for example, suggesting that the morpheme has a broader scope:

(13) oʜlan kitabɨ atmɨʃ
'the boy threw the book' - Theme, Patient

(42) oʜlan kɨza mɛktup jolamɨʃ
'the boy sent a letter to the girl' - Theme

And semantically, this matches the traditional concept of direct object or complement of the phrase.

Other cases – namely dative – are poorly motivated, because indirect objects/Recipients are marked using a separate set of endings indicating spatial relations. If indirect object (adjunct) marking can be explained via a system of spatial terms, then it is hard to justify its inclusion in a case system; and it would be an odd case system that lacks a dative. However, both object-marking and spatial morphemes occur in the same position on the noun (i.e. word-finally), and they are always mutually exclusive, suggesting if only weakly that they might constitute two separate aspects of a single case system.

For ease of glossing, I will refer to object marking as an accusative marker primarily indicating Patient in transitive sentences (but other potential thematic roles as well); and thematic roles such as Recipient, Goal, or Beneficiary in terms of the spatial relations they represent (e.g. TO, AT, and FROM), understanding that at the same time these can function as part of an elaborate case system.

The accusative is marked in Turkish with the morpheme {-V} where the vowel follows Harmony Rule #2. Just like with the POSSD marker, an epenthetic consonant /j/ is inserted before the morpheme in order to separate the vowels. This object marker, as mentioned earlier, is optional in certain contexts. It is much more common for the morpheme to appear on an object than not, but there are specific cases where it is more (or less) optional than others. Let's begin with the cases where object marking is most obligatory.

For starters, the morpheme is not typically omitted in sentences which are even mildly complex, such as those containing locative phrases. While we don't yet have the data to show this unequivocally, it only makes sense that, the more things which occur in a sentence, the greater the chance of confusion, and therefore the higher the likelihood that the object will be marked for clarity.[4] And our data is at least consistent with this claim, because no sentence of marginal complexity omits marking on the object.

Object marking is also obligatory when the same noun must take other marking as well. Consider the sentence in (92). Ordinarily, object-marking in simple Agent-Patient-Verb structures is optional. Yet here the additional morphology necessitates its presence, so that you (theoretically, as this has not yet been tested against the informant's judgments) will never see sentences of type (92B).

(92)	ajɨ-nɨn	jɛmɛʜ-in-i	ʧal-d-ɨn
	bear-POSSR	food-POSSD-ACC	steal-PAST-2SING
	'you stole the bear's food'		

(92B)	ajɨ-nɨn	jɛmɛʜ-in	ʧal-d-ɨn
	bear-POSSR	food-POSSD	steal-PAST-2SING

Finally, object marking is obligatory in non-typical sentence structures, such as fronting the direct or indirect object. Again, this is merely for a matter of clarity. Object marking is omitted in (47) and (48), but obligatory in (49) due to its unusual fronted position.

[4] In order to show that sentence complexity triggers obligatory marking, we would need a set of two sentences A and B, where A consists of a simple transitive verb phrase (e.g. 'the dog bit the boy'), and B includes both A and an additional phrase C (e.g. 'on the neck', 'at the party', 'last week', or a relative clause). If B was consistently marked for object but A were not, then we could safely assume that the object in B was marked for clarity in a complex sentence.

(47) oʜlan farɛ-jɛ kurabijɛ vɛr-miʃ
boy mouse-REC cookie give-PAST

(48) farɛ-jɛ oʜlan kurabijɛ vɛr-miʃ
mouse-REC boy cookie give-PAST

(49) kurabijɛ-ji oʜlan farɛ-jɛ vɛrmiʃ
cookie-ACC boy mouse-REC give-PAST

However, when the Patient is obvious from context and its position in the sentence is clear, object marking is clearly optional, as evident from the variation on *ɛlma* 'apple' below:

(31) kadɨn adam-a ɛlma vɛr-di
woman man-REC apple give-PAST
'the woman gave the man an apple'

(96) biz kɨz-a ɛlma-jɨ vɛr-d-ik
we girl-REC apple-ACC give-PAST-1PL
'we gave the girl an apple'

Object marking is also directly related to specificity. In addition to strong evidence from the data supporting a connection between object marking and definiteness, the informant also told us directly that the object marker is always used when talking about a specific referent, and omitted when the word may be referring to one of a number of objects, or it isn't specified. This phenomenon is especially noticeable on nouns ending in voiceless stops or affricates, as it elicits voicing on the consonant. Definiteness (and thus object marking) can also be triggered with demonstratives like 'this' or 'that'.

4.2.4 Pronouns

I classify pronouns here under nouns because they differ only mildly from other nouns, and follow the same patterns of morphology. The only difference is that they trigger slightly different agreement marking on both the noun (e.g. in possessive agreement) and the verb (subject agreement). Also pronouns don't pluralize in quite the same way as standard nouns (e.g. *sin* 'you' → *siniz* 'y'all; *o* 'he/she' → *onlar* 'they'). However, both these differences can be resolved in nouns and pronouns are seen as belonging to a larger, overarching system of marking and agreement, of which 'regular' nouns constitute only a specific class; that is, regular nouns take only third

person agreement, whereas other pronouns take the agreement for their respective persons. If we were to summarize that system, it would look like Table 5.

	Number	NOM./SUB.	ACC.	POST.
SING.	1	b+{ɛn} → bɛn	bɛn+{V_2} → bɛni	bɛn+{V_1} → bana
	2	s+{ɛn} → sɛn	sɛn+{V_2} → sɛni	sɛn+{V_1} → sana
	3	o	o+{V_2} → onu	o+{V_1} → ona
PL.	1	b+{Vz} → biz	biz+{V_2} → bizi	biz+{V_1} → bizɛ
	2	s+{Vz}+({lVr}) → siz(lɛr)	siz+{V_2} → sizi	siz+{V_1} → siz
	3	o+{lVr} → onlar	onlar+{V_2} → onlarɨ	onlar+{V_1} → onlara
		POSSR.	POSSD.	
SING.	1	bɛn+{(n)Vm} → bɛnim	{Vm}	
	2	sɛn+{(n)Vn} → sɛnim	{Vn}	
	3	o+{(n)Vn} → onun	{sV(n)}	
PL.	1	biz+{(n)Vm} → bizim	{Vm}+{Vz} → {VmVz}	
	2	siz+{(n)Vn} → sizin	{sVn}+{Vz} → {sVnVz}	
	3	onlar+{(n)Vn} → onlarɨn	{sV(n)}	

Table 5

First we need to iron out the details of this chart. The singular subject/nominative form appears to be the root from which all other forms of the pronoun are derived. The core element of the first, second, and third person are the phonemes /b/, /s/, and /o/ respectively, and it is from these basic elements that all other variations of the pronoun are formed. The nominative singular is formed by the addition of {ɛn} in the first and second person, while the nominative plurals are derived by affixing the so-called historical plural morpheme onto the singular (or rather, by changing the ending on the singular; it would seem these are suppletive forms); the exception is the third person, where the plural is formed by adding the nominal morpheme {-lVr}, which in this case causes an epenthetic /n/ to appear.

The accusative form of the pronoun is formed by adding a vowel following Harmony Rule #2 (indicated by the subscript on the vowel). This again causes epenthesis on the third singular. The dative/indirect object form is derived by the addition of a vowel following Harmony Rule #1.

The first person possessor morpheme is {-(n)Vm}, while the second and third person possessor morpheme is {-(n)Vn}. Clearly these two morphemes are related, with variation on the final consonant explained by agreement to the noun it is modifying. And while the /n/ at the beginning of the morpheme only appears on the third person when preceded by a vowel, it is reasonable (or at least consistent) to assume that the /n/ is always part of the morpheme, and by virtue of the rule {(n)Vn} → /Vn/ | C_ it deletes in the first and second person.

The POSSD morpheme in the first person is {Vm}, {Vn} on the second, and {sVn} on the third. In the third person, the /n/ occurs in word-medial positions and deletes word-finally. This form also appears to contain internal morphology, so that /s/ is the third person marker in entities being possessed. The presence of a nasal consonant in all three of these morphemes suggests that they are related historically. This three-way distinction between first, second, and third person on the POSSD morpheme allows the hearer to immediately infer an implied POSSR if one is not given (as it often is not). Finally, the plural POSSD morpheme is once again formed by adding the historical plural {-Vz} to the singular POSSD morpheme in the first and second person, and by adding the 'modern' plural {-lVr} in the third.

4.2.5 Case Marking

Despite reservations to the contrary, it does seem that the best way to analyze the variety of endings on the noun is as a case system. But first let's look at evidence against this claim.

Turkish contains a very fundamental three-way spatial distinction between TO, AT, and FROM, each marked with a separate morpheme ({-V}, {-Vn}, and {-dVn} respectively). Semantically, these are very much like postpositions, in that they all deal with spatial concepts, and are used to indicate locative relations, movement to or from an entity, and (as a metaphorical extension of the preceding) the Recipient or Beneficiary of a sentence. They have a very broad distribution and are used to construct a variety of different expressions.

Typologically, case markings combine primarily with nouns, while adpositions can combine with phrases of different categories, like in (129).

(129) øndʒɛ-dɛn koʃ-ujor-d-um
before-FROM run-PROG-PAST-1SING
'I was running earlier'

Adpositions also tend to be monomorphemic and non-inflecting, as ours are. Finally, Nihan consistently separates these postpositions from the rest of the word with an apostrophe, so that *øndʒɛdɛn* above would in fact be written *øndʒɛ'dɛn*, a minor orthographic point with strong implications.

Another morpheme very much like a postposition is the Instrument marker (INSTR), marked by the morpheme {-lV}, which takes Harmony Rule #1. Another compelling argument for analyzing these morphemes as postpositions is that they can be used in addition to (some) other case endings, namely the genitive/possessive. Case markers are typically mutually exclusive.

These ‘postpositions’ *are* mutually exclusive with object marking, however. They take vowel harmony with the noun, suggesting that they are in fact part of noun morphology and not a free morpheme. We have also seen no evidence to suggest that these morphemes can operate over an entire phrase (such as in English where one can say ‘I gave the book *to him and her*’). These morphemes appear inseparable from the noun. Lastly and most importantly, these ‘postpositions’ do serve the syntactic and semantic functions of object marking. It is probably a safe bet, however, that some of these case markings, as I will call them from here on out, and especially the locative morphemes were originally postpositions of some sort. This is fairly typical of verb-final/head-final languages.

What this means is that Turkish has a somewhat complex system of case marking, that is likely still developing. It would consist as follows:

Nominative	ø
Genitive	{-(n)Vn}
Accusative	{-V_1}
Dative/TO	{-V_2}
Locative/AT	{-dV}
Derivative/FROM	{-dVn}
Instrumental	{-lV}

No distinction is made between telic (e.g. ‘to’) and atelic (e.g. ‘towards’) relations. Instead, Turkish has a simple three-way spatial system distinguishing between static (AT) and directional (TO, FROM) concepts. In addition to their purely locative functions, these case endings have been given extended metaphorical uses, so that ‘TO’, for example, has extended its use to mark Recipients or Beneficiaries as well.

Nouns marked by Dative/TO are best seen as the Goal of the sentence, which includes both Recipients and many locatives. There is generally a sense of movement towards the entity. The concept is marked by the morpheme {-V}, where the vowel follows Harmony Rule #1. (43) is an example of a simple locative construction using TO.

(43) iki kadɨn køpɛʜ-i jɛr-ɛ koj-muʃ-lar
two woman dog-ACC ground-TO put-PAST-PL
‘two women put the dog on the ground’

Nouns marked by Derivative/FROM are best described as the Source of the sentence, including agents and other locatives. It is the opposite of TO, involving an agent and a sense of movement from an entity. It is marked by the morpheme {-dVn}, where the vowel follows Harmony Rule #2. (72) is an example of a locative construction

using FROM in a possessive phrase, although simple constructions like (43) above are also possible.

(72)	aiʃɛ	gøkhan-ɨn	kɛdi-sin-dɛn	kɨl	al-mɨʃ
	Ayşe	Gökhan-POSSR	cat-POSSD-FROM	hair	take-PAST
	‘Ayşe took hair from Gökhan’s cat’				

The Locative/AT case marking, unlike TO and FROM, involves no agent or recipient, but is entirely passive, and typically expresses being or location. The morpheme for this case marker is {-dV}, where the vowel takes Harmony Rule #1. In addition to the two types of constructions seen above, these case markings are used in all spatial relations between entities, but do so using a possessive phrase mirroring the structure of the example in (72).

(59)	kitap	masa-nɨn	yst-yn-dɛ
	book	table-POSSR	top-POSSD-AT
	‘the book is on the table’		

Literally, we can construe this phrase as ‘the book is at the top of the table’. This is how most locative constructions are formed. Whereas English prepositions encode meanings like ‘under’ and ‘on top of’ directly, in Turkish it must be done through a combination of a locative noun (e.g. ‘top’, ‘bottom’) and one of the three spatial case markings.

4.3 Verbs

Our data set contains 23 verbs, including verbs of motion, sensory verbs, transitive, instransitive, and ditransitive verbs, plus others. Like nouns, these subcategories all share basic characteristics which justify lumping them under the category ‘verb’. First, the verb is always the last word in the sentence, in every sentence in the data. We can thus call this the verb position and assume that all the words in this position share the common feature of ‘verb’ (this, of course, is subject to change as we introduce more data on relative clauses, infinitives, etc.). Semantically, all our elicited verbs fit the typical characterization of the category, relating information about either an action or a state of affairs. Finally, verbs take morphology for person, number, tense, and what I tentatively call aspect.

The structure of the Turkish verb is as follows:

STEM + NEGATION + PROGRESSIVE + TENSE + SUBJECT MARKER

Verb morphology again follows the two basic rules for vowel harmony, and undergoes similar epenthetic processes as does noun morphology. We will see how each of these works as we examine each type of morpheme on the verb. Also, the infinitive of the verb is formed by adding the morpheme {-mVk} to the stem, Harmony Rule #2, and the imperative (at least in the singular) is formed simply by using the stem without any additional morphology. Both these latter observations come from comments made by the informant, however, and not our elicitation data.

4.3.1 Negation

Negation on the verb is marked with the morpheme {-mV-}, immediately after the stem, following Harmony Rule #1. The vowel falls away before the progressive aspectual morpheme /Vjor/, as in (151) below.

(151) armut-u jɛ-m-ijor-lar
pear-ACC eat-NEG-PROG-3PL
'they are not eating the pear'

We know that the vowel between the /m/ and the /j/ belongs to the progressive marker, and not the negative marker, because it does not fit the vowel harmony rule for the negation marker. Instead, it fits Harmony Rule #2, which is the rule for the progressive marker. And because vowels in such position consistently match Rule #2, we can assume that the vowel on {-mV-} has undergone deletion, rather than some process of assimilation with the nearby vowel. This is one of only two cases in our data so far where vowels drop, rather than inserting an epenthetic consonant (the other situation being the vowel in the past tense marker {-dV-}).

Finally, a quick example showing the negation marker in action, without vowel deletion:

(133) jarɨn koʃ-maj-adʒaʜ-ɨm
tomorrow run-NEG-FUT-1SING
'I will not run tomorrow'

4.3.2 Aspect

The motivation for a system of aspect in Turkish verbs is minimal at best, but it seems that it is the best method of describing the progressive marker in Turkish. The progressive marker semantically is what we think of in English as the progressive, and is usually treated as an aspect in languages, rather than tense, because it describes how the action occurs (e.g. ongoing), not precisely when. Nor does it fit into the slot

for tense on the Turkish verb; instead, it coincides with the tense marker to give further information about how/when the action took place. Finally, there are inklings of aspect in the so-called 'hearsay' marker, which stands in opposition to the simple, completed past. This, however, does fit into our tense slot in the verb, and so rather than viewing it as aspect, it might be better seen as a sort of 'general' or 'indefinite' tense, thus giving us a tense system (based on the data so far) consisting of present, past, future, and indefinite/hearsay tenses, which can then be further modified by aspect (although currently the only aspectual marker we have is, as already mentioned, the progressive). Example (127) also proves that {-mVʃ-}, the habitual marker, must mark for tense, because it co-occurs with the progressive marker {-Vjor-}.

(127) dyn jɛmɛk j-ijor-muʃ
yesterday food eat-PROG-HAB
'he was eating yesterday'

The progressive morpheme is {-Vjor}, and takes Harmony Rule #2. The interesting thing about this morpheme is that it is the first we have encountered that has a static vowel, independent of vowel harmony. Additionally, this morpheme has the unusual ability to cause other morphemes to drop their vowels (namely the negation marker and future tense marker, but also at the end of some verb stems, such as the verb 'eat' in (127) above – normally realized as /jɛ/ – and 'kick' in (112) in the next section – normally realized as /tɛkmɛlɛ/), instead of triggering consonant epenthesis. These facts may be related, although it is unclear how.

The progressive marker can be used with the past, present, and habitual tenses; however the future tense marker subsumes both what English speakers would think of as the future and future progressive tenses (i.e. the future tense marker stands for both), meaning that {-Vjor-} would be, not just redundant, but in fact incorrect if used with the future tense, as shown by (156).

(156) *koʃ-ujor-ʤak
run-PROG-FUT-Ø (3SING)
'she will be running'

Note also that here the vowel at the front of the future morpheme (which is typically realized as {-Vʤak-}) drops away entirely.

4.3.3 Tense

The Turkish tense system – at least insofar as we have seen – is relatively simple, consisting of four tenses (simple/completed past, present, future, and habitual/hearsay/indefinite) and one aspect (progressive). Each of these tenses and aspect receive their own unique morpheme on the verb. Since aspect has already been dealt with, we'll begin with the present.

The present tense is unmarked on the verb (or perhaps marked by Ø), and has occurred in only two instances: present tense locative constructions, and present progressive verbs. Present tense locative constructions lack a verb entirely (or any form of copula), as shown in (59).

(59)	kitap	masa-nɨn	yst-yn-dɛ	Ø
	book	table-POSSR	top-POSSD-AT	PRES
	'the book is on the table'			

In sentences like (112) below, there is no morpheme in the tense slot on the verb – only the progressive marker {-Vjor-}. It may be that this progressive marker is the only way to indicate present tense on the verb. In fact, in eliciting what we might consider in English as the present indefinite, the verb retained the progressive marker (see (202) below), suggesting that Turkish has only one present tense – the present progressive – which encompasses both aspects of our English present tense. So far, the present tense has yet to occur on a verb without the co-occurrence of the progressive.

(112)	køpɛʜ-i	tɛkmɛl-ijor-um
	dog-ACC	kick-PROG-1SING
	'I am kicking the dog	

(202)	kɛndi-mi	sɛv-ijor-um
	RFLX-POSSD	love-PROG-1SING
	'I love myself'	

The completed past is marked with the morpheme {-dV-}, taking Harmony Rule #2. It conveys a simple or completed action in the past tense. The vowel drops before any subject agreement morpheme which begins in a vowel (i.e. all of them). The only reason we can be sure that the past morpheme contains a vowel is the subject agreement morphemes for third person plural – it is the only subject agreement marker which we are sure does not begin in a vowel (because it occurs on nouns and combines with other tenses, always in the same form, {-lVr}), and yet the vowel is still present in cases like (104) below. Even the third singular like (103) (below) cannot tell

us unequivocally that the past morpheme is {-dV-}, because it is quite possible that the marker for third singular is simply {-V} (thus making the past marker just the consonant {-d-}). However, reference to the third plural marker has shown us that the past marker is in fact {-dV-}, and thus the third singular marker must be {-Ø}.

(103) o biz-i tɛkmɛlɛ-di-Ø
he us-ACC kick-PAST-3SING
'he kicked us'

(104) onlar jɛmɛʜ-i biz-ɛ vɛr-di-lɛr
they food-ACC us-TO give-PAST-3PL
'they gave food to us'

The future tense is marked by the morpheme {-VʤVʜ-}, which has some interesting phonological features. To begin with, the two vowel do not follow the same harmony rules. The first vowel always takes Harmony Rule #2, and so we would expect the second vowel to always be the same (if it also followed the same harmony rule). But the second vowel is always /a/ or /ɛ/, depending on the preceding vowel, meaning that it follows Harmony Rule #1. The implications of this fact are unclear. It might mean that the future morpheme has some sort of (historical) internal morphology; or it might simply be for simplicity's sake (i.e. simplicity in processing the phonology of the word). What I mean by this is that, while Rule #2 depends on both roundness and frontness for determining harmony, Rule #1 depends only on frontness. Therefore it might be that it is psychologically simpler for a speaker to process the phonological rules in this order when two vowels must be accounted for. Unfortunately, the only other morpheme with two internal vowels that we have extensive data on is {-Vjor-}, and the second vowel is entirely stable. Still, what simpler way of determing the second vowel in a morpheme than for it to be immutable? Another potential morpheme following a similar process as {-VʤVʜ-} might be the tense marker {-VbVlVr-} 'might', but as yet there is not enough data to determine if it functions in the same way as {-VʤVk-}. In any case, this is highly speculative – merely an observation that there might be some pattern determining successive vowels in a morpheme.

Like our observations with nouns earlier, the /ʜ/ of the future morpheme becomes devoiced finally, from /ʜ/ → /k/, but its voiced quality appears when combined with basic subject agreement morphology, as in (133):

(133) jarɨn koʃ-maj-aʤaʜ-ɨm
tomorrow run-NEG-FUT-1SING
'I will not run tomorrow'

Nihan once described the tense morpheme {-mVʃ-} as a 'habitual past', however I have difficulty motivating this based on the data. It is much easier and more consistent with the data to analyze the morpheme in the way that Nihan originally described to us – as a 'hearsay' marker; that is, the morpheme is used on verbs where the speaker did not directly witness the action. In a sense, this is at least consistent with calling it an indefinite marker, as indefinite statements (e.g. 'I go to school') can't really be witnessed if they refer to some type of ongoing or indefinite action. So it is in this sense, perhaps, that the morpheme {-mVʃ-} might be considered habitual, or ongoing. However I will continue referring to the morpheme as the hearsay tense, as it is the simplest analysis.

The hearsay morpheme {-mVʃ-} takes Harmony Rule #2, and is entirely regular in its occurrence (i.e. it undergoes no phonological process aside from vowel harmony, at least in as far as we are aware). It can occur with the past and present tenses, but not the future. This can be explained semantically if we realize that it seems odd to discuss witnessing or not witnessing an action which has yet to occur. For similar reasons, the hearsay morpheme occurs rather infrequently with interrogatives, and so far only with the interrogative 'where'. When it does occur with interrogatives, as in (190), it must mean something other than the simple past tense, as indicated in the translation below.

(190) køpɛk nɛrɛ-jɛ koj-muʃ-lar
dog WH-TO put-HEAR-3PL
'where did you hear that they put the dog?'

A less ambiguous translation might be 'where did they put the dog, that you heard of?' So to summarize, the {-mVʃ-} morpheme is seen most easily as a hearsay marker, but with ample room that it may in fact be some type of indefinite or habitual marker.

4.3.4 Subject Agreement

The subject agreement suffixes for Turkish verbs are given in Table 6 below. All the vowels take Harmony Rule #2 with the exception of the third plural. They are relatively

	SING	PL
1	-Vm	-Vk (FUT: -Vz)
2	-(s)Vn	-(s)VnVz
3	Ø	-lVr

Table 6

straightforward, the only variation occurring in said vowel harmony, a phonological

process on the markers for second person, and a future tense variant of the first person plural.

As stated previously, the initial vowels on these morphemes trigger vowel deletion on the past morpheme {-dV-}, and elicit voicing of the final consonant in the future morpheme {-VʤVʜ-}. The second person markers are an exception to this latter fact. In the majority of cases, the second person markers appear without the initial /s/, as in (92) and (99) below.

(92) ajɨ-nɨn jɛmɛʜ-in-i ʧal-d-ɨn
bear-POSSR food-POSSD-ACC steal-PAST-2SING
'you stole the bear's food'

(99) siz aiʃɛ-nin jɛmɛʜ-in-i jɛ-d-iniz
you Ayşe-POSSR food-POSSD-ACC eat-PAST-2PL
'y'all ate Ayşe's food'

Thus one would expect them to trigger intervocalic voicing when preceded by /Vk/, but in fact constructions like (134) are the case.

(134) jarɨn armut-u jij-ɛʤɛk-sin
tomorrow pear-ACC eat-FUT-2SING
'you will eat the pear tomorrow'

Interestingly, this is not the only place this phenomenon occurs:

(118) sɛn koʃ-ujor-sun
you run-PROG-2SING
'you are running'

(140) ʃɛrap-ɨ øbyrs-y gyn iʧ-ɛbilir-sin
wine-ACC tomorrow-TO day drink-might-2SING
'you might drink wine the day after tomorrow'

We have very little data on the {-ɛbilir-} morpheme, however we do see instances of vowel deletion after /r/ in other places (cf. (187) and (188) below). This suggests the possibility that there may be a deleted vowel in the examples above. It seems unwise to guess that the /s/ might be the result of epenthesis, because consonant epenthesis typically occurs between other consonants, and that is not the case here. Whatever the explanation, the pattern is clear – the /s/ on the second person morphemes appears only after an /r/.

(187) kitap nɛrdɛ
OR
(188) kitap nɛrɛdɛ
'where is the book?'

So the fact that the second person morphemes do not trigger voicing on {-VʤVʜ-} suggests that the /s/ is part of the second person morphemes. Clearly the /s/ is psychologically salient as part of the second person markers, enough that it 'overrides', so to speak, the voicing that might otherwise be caused by the vowel. The tricky part is explaining why a consonant disappears from the morpheme when when it follows a vowel. Typically, Turkish's tendency to avoid vowel clusters would motivate the presence of /s/; instead the /s/ is not only dropped but another vowel must be deleted off the end of the preceding tense morpheme in order to adhere to Turkish syllable structure (CVC). There is nothing in the data to explain why such a complex process is preferred over the simpler one of maintaining the /s/ on the second person morphemes.

Turning to the variation on the first person plural form, it is worth noting that the future variant of the morpheme ({-Vz}) resembles the second plural marker {-VnVz}. This similarity suggests that historically there was a plural marker – probably {-Vz} – which may have been suffixed to the singular person agreement morphemes (perhaps a rule like {singular agreement marker} + {-Vz} → {plural agreement marker}). There is additional evidence for internal morphology on the second plural morpheme as well. First, the 'root' of the morpheme is the same ast he second singular morpheme (i.e. {-Vn}. Second, the second singular has the same phonological rules regarding /s/ deletion as the plural, again suggesting that the basic morphology of these two morphemes is the same.

The last piece of data which needs explaining is why the third person plural marker takes a different harmony rule than the other subject agreement morphemes. The most likely explanation is that {-lVr} is not the original plural marker, that it merely emerged via analogy to nominal plural marking. This is consistent with the hypothesis that there was originally a different plural marker {-Vz}, which came to be replaced on the third plural by {-lVr}, via analogy to noun morphology.

4.4 Modifiers

This section covers adjectives, determiners, adverbs, and quantifiers, based on the observation that all three serve to modify nouns (though they may be at different levels or have different functions syntactically) , and precede the nouns they modify. Additionally, there is little data to give a fleshed-out analysis of these three categories.

For the moment the best thing to do is simply note how the three are different, justifying their placement into three different (sub?)categories.

Adjectives always immediately preced the noun they modify, and take no agreement or other morphology. Several adjectives (at least two) can be strung together, as in (64).

(64) kytʃyk kahvɛrɛngi køpɛk masa-nɨn alt-ɨn-da otur-du-Ø
small brown dog table-POSSR bottom-POSSD-AT sit-PAST-3SING
'the small brown dog sat under the table'

Adjectives don't inflect for plural, but rather plurality is marked simply on the noun. Conversely, quantifiers such as numbers do *not* allow for plural marking on the noun (compare the phrases taken from (55), (34) and (35)). The exception to this rule is when the plural noun takes certain case markings, as in (225).

(55) byk aj**ɨ**-lar
big bear-PL
'big bears'

(34) dørt armut
four pear
'four pears'

(35) *dørt armut-lar
four pear-PL
'four pears'

(225) on adam-lar-la
ten man-PL-INSTR
'along with ten men'

We do not yet have data on the order of quantifiers and adjectives put together, but hazarding a guess, I imagine quantifiers precede adjectives in a noun phrase. Demonstratives like *o* or *bu* we know always precede adjectives in a noun phrase, and do allow plural morphology on the noun phrase (probably because the determiners themselves do not inflect for plurality). While adjectives can be modified by other adjectives or demonstratives, demonstratives cannot, and there is no data as to whether or not quantifiers may also take additional modifiers such as demonstratives.

We have also encountered one possible determiner, *bir* 'one', but which is best seen as merely another quantifier.

These morphological and syntactic differences are enough to motivate at least different subcategories of modifiers, and more likely different categories entirely, based simply on the semantic notions of the words. Adjectives serve to directly modify a noun or other adjective, determiners serve to indicate specificity or definiteness on the noun, demonstratives point out relative locations, and quantifiers serve to specify the number of the noun.

Finally, there seems ample evidence for the category of adverb. While most 'adverbial' phrases we have seen so far might be explained by locative or nominal constructions like (130), a number (like the presence of words for 'yesterday' and 'more' in other examples (namely (127) and (128)) cannot.

(130) øndʒɛ-dɛn koʃ-d-um
before-FROM run-PAST-1SING
'I ran earlier'

4.5 Expectations & Universals

I have tried to motivate each lexical category as it is considered over the course of the analysis, using both facts from the data (i.e. morphology, syntax, semantics, etc.) and linguistic universals. However it still serves us to compare that analysis to what we should expect typologically.

All languages distinguish between nouns and verbs, with a continuum of other categories in between (Whaley, 1997). Nouns typically express thematic roles like Agent, Patient, or Recipient, can be modified by determiners, adjectives, or other modifiers, and appear in both subject and object positions – all of which Turkish nouns do. Morphologically, nouns tend to take morphemes marking plural, possession, or case. Turkish marks plural on most of its nouns, clearly marks possession on both the noun and modifying noun (i.e. POSSD and POSSR), and seems to have some type of case-marking system. Semantically, they indicate entities or referents, most often persons, places, or things.

Verbs are generally marked for person, number, tense, voice, mood, aspect and other features. In Turkish, we know of the existence of verbal morphemes to mark tense, number, aspect (maybe), and mood (so far just indicative, imperative, and infinitive). Some languages have verbs which encode pragmatic information about the speaker, as in Turkish, where there is a morpheme marking how the verb relates to the speaker's knowledge (the hearsay morpheme). Typologically, we should expect a difference between active and passive voice, but this has yet to be elicited that we know of.

4.6 Allomorphs

Typically, word-final stops become devoiced in Turkish. With the addition of the accusative marker (or any morpheme starting with a vowel), the stop stays voiced between vowels, so that /k/ → /ɰ/, /p/ → /b/, /ʧ/ → /ʤ/, and presumably /t/ → /d/, although this latter case has not yet occurred in our data. With object marking, this helps make the difference in specificity overt, as in (1) and (6):

(1)	kɨz	kitab-ɨ	oku-du
	girl	book-ACC	read-PAST
	'the girl read the book'		

(6)	kɨz	bir	kitap	oku-du
	girl	one	book	read-PAST
	'the girl read a book'			

In the first and second singular personal pronouns, the addition of the Dative morpheme triggers the irregular forms *bana* and *sana* ('to me' and 'to you' respectively). This vowel change on the stem is likely explained by the fact that /ɛ/ is realized phonetically as [a] before a nasal ([ɛ] → [a] | _č), so vowel harmony and the affixation of another /a/ makes this phonetic realization overt. The third singular again takes an epenthetic /n/, and for some reason the vowel on the second plural doesn't become realized. However, this datum may simply be erroneous, because portions of Table 5 were not elicited, but given by Nihan directly, so there is no point of reference for comparison.

While the possessive morphemes also undergo considerable morphological variation, there is decent evidence to suggest that they all stem from the same underlying morpheme. The possessive morpheme (whether POSSR or POSSD) can be realized via the following allomorphs:

-Vn	POSSR after consonant
-nVn	POSSR after vowel
-Vn	POSSD for second person, and third after consonants
-sVn	POSSD for third person
-Vm	POSSD for first person

These morphemes seem to have an internal morphology, though the exact pattern is hard to discern. But to point out the basics, -Vm and -sVn look to have first and third person agreement respectively (the /m/ and /s/), so we imagine that -Vn too is in some way taking second person agreement (perhaps with the /n/). The /s/ in the third

person POSSD marker appears only when preceded by a vowel (to avoid vowel clusters), just as the /n/ in the POSSR morpheme appears only when preceded by a vowel. The implication here is that there is one underlying morpheme for possessives, with perhaps four allomorphs: one for POSSR {-(n)Vn}, and three for POSSD, which must agree with the person of the POSSR. Those three are {-Vm}, {-Vn}, and {-(s)Vn} for first, second, and third persons respectively.

4.7 Problems & Exceptions

An initial problem in determining the rules for vowel harmony on the morphemes whether whether the vowel in the morpheme harmonizes with the root of the word or the immediately preceding vowel. Unfortunately, cases like *mantar-lar-ɨn-ɨ*, *kaja-lar-ɨn*, *marɛ-lɛr-in*, and *kadɨn-lar-a* are all ambiguous as to which vowels harmonize with which. Fortunately one example – *on-lar-ɨ* – shows that the vowel /ɨ/ must harmonize with the vowel immediately preceding it, and not the stem. Later data (not included in this analysis) also confirms this, and this fits what we should expect typologically concerning vowel harmony.

There is only one unexplained segment in the data, and it occurs only twice, in (56) and (162), which are given below along with (129) for comparison.

(56) bu marɛ-dɛ-ki kɨllɨ ajɨ o kaja-nɨn yst-yn-dɛ uju-muʃ
DEM cave-AT-? hairy bear DEM rock-POSSR top-POSSD-AT sleep-HEAR
'the hairy bear slept in this cave on those rocks'

(162) jarɨn-dan øndʒɛ-ki gyn
tomorrow-FROM before-? day
'the day before tomorrow'

(129) øndʒɛ-dɛn koʃ-ujor-d-um
before-FROM run-PROG-PAST-1SING
'I was running earlier'

The mystery morpheme is {-ki}, or possibly {-Vki}, although (129) suggests that the vowel belongs to the 'before' morpheme. It is unclear precisely what {-ki} is. Our later data shows that it functions as a complementizer, and that in itself is certainly a satisfactory explanation. But personal data from my project more closely mirrors the use give above, where instead a locative phrase is used to further specify details about a situation. In the same way, a complement phrase could be said to be further specifying details about the situation.

§5. Syntactic Analysis

5.1 Phrase Structure Rules

Here we use the same set of 300 elicited sentences to analyze the syntax of Turkish. The Turkish language is primarily head-final, and as is common for head-final languages, follows SOV word order. There are, however, numerous exceptions to this, which will be discussed later.

There are six primary types of phrases in Turkish, namely the noun phrase (NP), verb phrase (VP), adjective/adverb phrase (AP), determiner phrase (DP), complementizer phrase (CP), and relative clause (RC), although it may be possible to view this as simply another NP or even a sentence (S). Finally, a sentence (S) is formed using an NP and VP. The phrase structure rule (PSR) for the level of the sentence is generally S → NP VP, although later we will examine how sometimes elements of the VP can be fronted to precede the NP.

The PSR for NPs is as follows: [5]

NP → (DP) (QU) (NP/RC) (AP) N

With examples given below:

(57) kɨllɨ ajɨ-lar
hairy bear-PL
'the hairy bears'

(251) mavi ɛlbisɛ-ji gij-ɛn kɨz
blue dress-OBJ wear-REL girl
'the girl who is wearing the blue dress'

(80) kɛdi-nin kahvɛrɛngi jɛmɛʜ-in-i
cat-POSSR brown food-POSSD-OBJ
'the cat's brown food'

The example in (57) is a noun modified by an AP, while in (251) the noun is modified by a RC which itself contains an AP. In (80) the noun is modified by both an AP and an NP, where the NP is the possessive 'cat's'. These possessive NPs are the

[5] A parenthesis indicates that the enclosed phrase is optional, while an asterisk indicates that the phrase may be used multiple times consecutively. A slash indicates a choice between two phrases, so that their use is mutually exclusive.

primary motivation for the inclusion of a recursive NP in the PSR above. Consider the example in (78), where the recursive NP is bracketed:

78) [kahvɛrɛngi kɛdi-nin] jɛmɛʜ-in-i
brown cat-POSSR food-POSSD-OBJ
'the [brown cat's] food'

So nouns can modify other possessive nouns. It might also be possible to interpet the possessive marker as a determiner, so that we would view possessives as nouns being modified by a DP. As our NP PSR already includes an optional DP, this would allow us to drop the recursional NP entirely.

As far as our data shows, RCs and recursive NPs do not co-occur, hence the choice of (NP/RC). At the same time, a possessed RC (e.g. 'the boy who went to school's dog) was never elicited, and we might imagine that in Turkish it is possible to say something equivalent to the English.

Finally, note that the noun is phrase final, with its modifier/adjunct immediately preceding.

The VP and CP are formed as follows:

VP → (AP) (NP)* (CP) V (CP)
CP → (C) S

The presence of a complementizer in the CP is still open to debate, as the status of the *ki* morpheme which might serve as a complementizer is still undecided. Several examples of VPs are given below. Note that a CP, when used with the *ki* morpheme, follows the verb (although this is less natural), and precedes the verb without it, as in x and (290).

(118) koʃujorsun
run-PROG-2SING
'you are running'

(126) dyn ji-jor-d-u
yesterday eat-PROG-PAST-3SING

OR dyn jɛmɛk ji-jor-d-u
yesterday food eat-PROG-PAST-3SING

(147) kitab-ɨ køpɛk-lɛr-ɛ vɛr-mɛ-d-i
book-OBJ dog-PL-REC give-NEG-PAST-3SING
'he did not give the book to the dog'

(294) kork-ujor-um-ki onlar git-di-lɛr
fear-PROG-1S-COMP they go-PAST-3PL
'I fear that they left'

(290) tɛrkɛt-di-ʜin-dɛn kork-ujor-lar
leave-PAST-POSSD-FROM fear-PROG-3PL
'they fear that you left'

It is important to note that a great deal is conflated into the single element (NP)* in the VP PSR. To begin with, the NP is only made optional by intransitive verbs. For transitive sentences, the PSR would be VP → (AP) (NP) NP (NP) V, where the obligatory NP is the direct object, and the optional NPs are either the indirect object or a locative construction, where generally the indirect object precedes the direct object, and the locative follows it (see 40 and 60 below), making it necessary to have an optional NP to either side of the direct object.

(40) køpɛʜ-ɛ kitab-ɨ vɛr-miʃ
dog-TO book-OBJ give-HEAR
'he gave the book to the dog'

(60) kitab-ɨ masa-nɨn yst-yn-ɛ koj-d-u
book-OBJ table-POSSR top-POSSD-TO put-PAST-3SING
'he put the book on the table'

Notice, however, that (147) above breaks this pattern, and places the indirect object after the object. This is decidedly the marked version, because it forces object marking on the direct object, whereas direct objects which immediately precede the verb sometimes lack object marking. Also, it is most common in Turkish for the complement to immediately precede the head of the phrase.

The NP in the VP PSR also includes instrumental constructions, such as (220), and RC constructions such as (231) (although many of these RCs are themselves just an example of possessive NPs, including (231) below).

(220) kaʃɨk-la jɛ-di-m
spoon-INSTR eat-PAST-1SING
'I ate with a spoon'

(231) bɛn-im piʃir-di-ʜim øʜyn-y jɛ-d-i
I-POSSR cook-PAST-POSSD meal-OBJ eat-PAST-3SING
‘he ate the meal that I cooked’

Example (127) shows the motivation for the inclusion of the AP before the NP:

(127) dyn jɛmɛk ji-jor-muʃ
yesterday food eat-PROG-HEAR
'he was eating yesterday'

The PSR for APs is:

AP → (AP) A

While it is quite possible that multiple APs can modify an adjective, we only have an example of there being one extra adjective (for a total of two), show in (64) below. It is reasonable to assume, however, that there can indeed be multiple APs modifying an adjective. Also note that adjectives require no special morphology in order to mark their function, and always immediately precede the noun they modify (i.e. they are always the adjunct of the noun).

(64) kytʃyk kahvɛrɛngi køpɛk
small brown dog
‘small brown dog’

Next, RCs are formed in the following way:

RC → (NP) VP (NP)

Some examples of RCs in use:

(241) kɨz-a kitap vɛr-ɛn oʜlan
girl-OBJ book give-REL boy
‘the boy who gave the girl a book’

(243) sokak-da koʃ-an oʜlan-ɨ
street-AT run-REL boy-OBJ
‘the boy who was running down the street’

(231)	bɛn-im	piʃir-di-ʜim	øʜyn-y	jɛ-d-i
	I-POSSR	cook-PAST-POSSD	meal-OBJ	eat-PAST-3SING
	'he ate the meal that I cooked'			

(237)	piʃir-ʤɛʜ-in	pasta-jɨ	ist-ijor-um
	cook-FUT-POSSD	cake-OBJ	want-PROG-1SING
	'I want the cake you will cook'		

The VP in (231) and (237), since it utilizes possessive morphology, might also be interpreted as a verbal noun, in which case there's no real justification for considering RCs to be anything other than a standard NP. However, the relative marking on the verbs in (241) and (243) is still suspect as being nominal. We shouldn't think that the vowel-harmonized morpheme {-Vn} is some type of possessor, because it lacks a corresponding possessed noun, and because it does not allow for other verbal morphology in the same way as the verbs in (231) and (237). Instead, the {-Vn} stands alone as the only morphology on the verb. Thus it seems that {-Vn} is a type of relativizer, necessitating the existence of the RC, instead of simply calling it another type of NP. Note, however, that we cannot be sure if the {-Vn} morpheme creates another nominal verb analogous to (231) and (237), or if the word remains verbal – we only know that some RCs (those in which the relativized noun is the subject of the RC) have their own morphology, motivating their placement into their own category of RC. And if we are to be strict about defining our categories based on morphological analogy, we will not even include the nominal verbs of the type in (231) and (237) into the category RC, instead calling them instances of possessive NPs (which just happen to serve the same function as what we call relative clauses in English).

All that aside, relative clauses do look remarkably like regular Turkish sentences, consisting of a subject, object, indirect object, and verb (with various parts being optional or part of the main clause). They also follow standard SOV word order as much as possible (variation being due to the fact that one or another part of the sentence must fulfill its position in the main clause). The objects of the relative clause also take standard object case marking, as in (237). This all suggests that relative clauses may best be represented in the PSRs as simply S, and it is a matter of morphology that relativized sentences take special possessive morphemes. An alternative analysis would be to say that the relativized sentence is primarily nominal, and to subsume it under the recursive NP.

(Two revisioned versions of the NP PSR – relative clauses now require no special category:)

NP → (DP) (QU) (NP/S) (AP) N
NP → (DP) (QU) (NP) (AP) N

When the relativized noun is part of a locative phrase, the possessive morphology is also used, but this is due to the fact that locatives in Turkish are not formed with prepositions per se, but rather by using possessive morphology. The inclusion of a second NP in the PSR allows for the nouns used to create prepositional/spatial concepts and other relationships. In (255), however, the verb and the spatial noun are split by the NP of the main clause. Interestingly, while this is exactly where the word would belong if it were part of the relativized clause, we know it belongs to the main clause instead because it lacks the marking for possessor.

(255)	kirli	bulaʃɨk-lar-ɨ	koj-duʜ-un	masa	yst-yn-y
	dirty	dish-PL-OBJ	put-PAST-POSSR	table	top-POSSD-TO

'the table that you put the dirty dishes on'

Finally, DPs consist simply of a determiner (D).

5.2 Word-Order Typology

The most common word orders for language are SVO and SOV, and Turkish is primarily an SOV language. SOV languages tend to be head-final, and Turkish, as we have seen, generally follows this model, as evidenced by the fact that the object precedes the verb in every sentence we have. In fact, Turkish serves as an excellent example of a head-final language. In each PSR, the head of the phrase comes last; the morphology for every syntactic category also comes at the end of the word. In verb-final languages the possessor almost always precedes the possessed, as is the case with Turkish. Verb-final languages also tend towards postpositions (which in Turkish have become realized as locative case marking), and a pattern of modifiers preceding the element modified. Turkish follows this latter pattern fairly consistently, certain relative and complementizing constructions being the notable exceptions.

There are, however, instances where the object precedes the subject in Turkish – a highly unusual phenomenon in languages – such as in (48) and (49). There is an overwhelming tendency in languages for the subject to preced the object.

Notice that, compared to (47), (49) is especially marked as the indirect object.

(47)	oʜlan	farɛ-jɛ	kurabijɛ	vɛr-miʃ
	boy	mouse-TO	cookie	give-HEAR

'the boy gave the mouse a cookie'

(48)	farɛ-jɛ	oʜlan	kurabijɛ	vɛr-miʃ	
	mouse-TO	boy	cookie	give-HEAR	
	'the boy gave the mouse a cookie'				

(49)	kurabijɛ-ji	oʜlan	farɛ-jɛ	vɛr-miʃ	
	cookie-OBJ	boy	mouse-TO	give-HEAR	
	'the boy gave the mouse a cookie'				

There are other instances of variation in word order as well, such as VPs split by the subject NP as we will see later on, and particular freedom to move the object NPs. This is likely due to the fact that case marking tends to make word order freer (Whaley, p. 81).

Despite variation in the order of the subject and the objects, Turkish obeys Greenberg's Universal 1, that subjects almost always precede the object – this is the predominant pattern in Turkish. However, it has a strong tendency toward violating the principle that the verb and the object are generally adjacent (Whaley, p. 84), particularly when locative phrases are used:

(60)	adam	kitab-ɨ	masa-nɨn	yst-yn-ɛ		koj-du
	man	book-OBJ	table-POSSR	top-POSSD-TO		put-PAST
	'the man put the book on the table'					

5.3 Word-Order Variation

Variation in word-order is the greatest in the VP. As already seen from the examples, and as illustrated specifically in (96), the direct and indirect objects can be easily switched, usually forcing marking on the direct object.

(96)	biz	ɛlma-jɨ	kɨz-a	vɛr-di-k	
	we	apple-OBJ	girl-TO	give-PAST-1PL	
	'we gave the apple to the girl'				

Again, this is likely caused by the flexibility given the language by case-marking, as evidenced by the fact that, in such nonstandard constructions, all objects must be marked for case.

As in (60) above, locative phrases have a strong tendency to separate the direct object from the verb. The generative tradition would probably attribute this to movement of some sort, but as the locative phrase is marked with the TO particle, it seems that the thematic relations of the sentence will still be psychologically salient to the speaker. Thus, in (60), even without the verb we are able to infer that the book

moves ‘to the tabletop’ in some manner. Not all locatives split the verb and its complement though:

(94) tabaʜ-ɨn-dan jɛmɛk ʧal-mɨʃ
plate-POSSD-FROM food steal-HEAR
‘he$_i$ stole food from his$_j$ plate’

There is also flexibility in the position of the subject NP in relation to RCs. Consider the examples below:

(232) køpɛʜ-in ɨsɨr-dɨ-ʜɨ kadɨn-a adam armut-u vɛr-di
dog-POSSR bite-PAST-POSSD woman-TO man pear-OBJ give-PAST

OR adam køpɛʜ-in ɨsɨr-dɨ-ʜɨ kadɨn-a armut-u vɛr-di
man dog-POSSR bite-PAST-POSSD woman-TO pear-OBJ give-PAST
‘the man gave a pear to the woman who the dog bit’

In the first version, it might seem like the subject NP has been haphazardly embedded into the VP, but in fact this is just another example of fronting the indirect object, except that in this case the object is modified by an RC. Most interestingly, this fronted version seems to be the preferred word-order for sentences containing RCs. In (234), for instance, the direct object is fronted, and while we can assume that an alternate version exists where *kadɨn* comes first in the sentence, this was the more ‘natural’ form we elicited.

(234) oʜlan-ɨn sɛv-di-ʜi køpɛʜ-i kadɨn ʧal-ɨjor
boy-POSSR like-PAST-POSSD dog-OBJ woman steal-PROG
‘the woman stole the dog which the boy likes’

So it seems that a variety of elements can be fronted in the Turkish sentence, while the rest of the sentence continues to obey predicted word-order. In (56), for instance, the event is being described using two locatives, one of which is fronted.

(56) bu marɛ-dɛ-ki kɨllɨ ajɨ o kaja-nɨn yst-yn-dɛ uju-muʃ
DET cave-AT-? hairy bear DET rock-POSSR top-POSSD-AT sleep-HEAR
‘the hairy bear slept in this cave on those rocks’

And we have already seen instances of fronted objects, suggesting perhaps an empty category at the front of any Turkish sentence which allows for movement, which is normally marked (in the case by *-ki*).

Generally, the sum collection of all this variation might suggest that there is less evidence for the category of VP than originally thought. This would mean instead that a sentence consists of a flat structure (a verb and several NPs). This does, however, fly in the face of most of the generative tradition.

5.4 Thematic Relations

The thematic roles in any given sentence describe the role that the argument plays with respect to the predicate (Carnie, 2002, p. 164). Often this information about the relationship between the argument and the verb is encoded in the morphology of a language, and this is the case with Turkish. We will briefly list the various thematic roles and how they are marked.

The agent in Turkish is associated with two syntactic positions: either the subject NP in a sentence, or the possessor NP in some types of RCs (see below). As the subject of a sentence, the agent is never marked. In RCs, the agent is the possessor of the relativized verb (i.e. the one doing the action of, or being the subject of, the relative clause). This fits with the semantic definition of an agent, which is typically the initiator of some action.

(1) kɨz kitab-ɨ oku-du
girl book-OBJ read-PAST
'the girl read the book'

(226) oʜlan-ɨn oku-du-ʜu kitab-ɨ ʧal-dɨ-m
boy-POSSR read-PAST-POSSD book-OBJ steal-PAST-1SING
'I stole the book which the boy read'

Finally, in certain RCs a word may play both a thematic role in the VP and play the agent of the relativized clause:

(240) fil-i ji-jɛn aslan-ɨ dojur-du-n
elephant-OBJ eat-REL lion-OBJ feed-PAST-2SING
'you fed the lion who ate the elephant'

Related to this is the experiencer role, which, based on our data, looks to have no grammatical distinction from the role of agent – both can serve as the subject of a sentence or relative clause, and are unmarked as the subject of the sentence. While the role of agent implies volition, the experiencer by definition is an entity which feels, perceives, or experiences events without volition. So Turkish does not mark for volition of action in its subjects.

The patient in a sentence undergoes the effect of some action, often with some change of state. In Turkish, the patient is often given the object marker, such as both ‘elephant’ and ‘lion’ in (240) above. As mentioned previously, the marker is optional depending on specificity. A very similar role is that of theme, which indicates an entity which is moved by an action, or whose location is described. As implied by the definition, this occurs most often with locatives:

(60)	adam	kitab-ɨ	masa-nɨn	yst-yn-ɛ	koj-du
	man	book-OBJ	table-POSSR	top-POSSD-TO	put-PAST
	‘the man put the book on the table’				

In this example, *kitabɨ* might be described as both the patient and the theme.

These thematic correlations vary greatly, however. The subject of a sentence can often be the theme, as in (6) (taken from the project data), or Experiencer as in (22).

(6)	para	ʤyzdan-ɨn	itʃ-in-dɛ
	money	wallet-POSSR	inside-POSSD-AT
	Theme	Location	
	‘the money is inside the wallet’		

(22)	oʜlan	gɛjiʜ-i	gør-myʃ
	boy	deer-OBJ	see-HEAR
	Experiencer	Patient	
	‘the boy saw the deer’		

Instrument, or the means by which an action is performed, is given its own case marker:

(220)	kaʃɨk-la	jɛ-di-m
	spoon-INSTR	eat-PAST-1SING
	‘I ate with a spoon’	

Turkish lacks overt marking for location, instead utilizing a possessive construction, as we will see later. We also have no available data on how the beneficiary, or entity for whose benefit the action was performed, is marked in Turkish.

Turkish does, however, explicitly mark for goal, source, and recipient. The source is the entity from which something moves, either literally or metaphorically, and in Turkish is marked by the morpheme {-dVn}. In my transcriptions, this morpheme has typically been translated as the concept FROM. It can be used to further specify the movement involved in a verb, as in (62), or physical movement, as in (72).

(62) kitap masa-nɨn yst-yn-dɛn dyʃ-dy
book table-POSSR top-POSSD-FROM fall-PAST
'the book fell off the table'

(72) ayʃɛ gøkhan-(n)ɨn kɛdi-sin-dɛn kɨl al-mɨʃ
Ayşe Gökhan-POSSR cat-POSSD-FROM hair take-HEAR
'Ayşe took hair from Gökhan's cat'

A goal is the entity towards which something moves, either literally or metaphorically, and the recipient is a specific type of goal involved in actions describing changes of possession. The goal and recipient are conflated in Turkish, and both are marked with the {-V} morpheme, translated here as TO. It can be used with change of possession, as in (39), and also location, as in (60).

39) oʜlan køpɛʜ-ɛ kitap vɛr-miʃ
boy dog-TO book give-HEAR
'the boy gave the book to the dog'

(60) adam kitab-ɨ masa-nɨn yst-yn-ɛ koj-du
man book-OBJ table-POSSR top-POSSD-TO put-PAST
'the man put the book on the table'

5.5 Some Specific Constructions

5.5.1 Question Formation

The way Turkish forms questions is remarkably simple – the sentences undergo no movement, the interrogatives take no special marking, and yes/no questions are marked using a simple morpheme on the verb.

In yes/no questions such as (163), the morpheme {-mV} is suffixed to the verb following the person marking. After the 2nd singular and 1st plural persons, however, the interrogative morpheme is affixed just before the person marking, but after the tense marker. (167) and (173) show examples of this. It is hard to say why these specific forms would be different – a best guess would simply be that the phonology of Turkish pushes more towards this reverse construction in these cases (or perhaps avoids the more standard construction, possibly to avoid consonant clusters).

(163) oʜlan kɨz-ɨ at-dɨ-mɨ
boy girl-OBJ throw-PAST-Q?
'did the boy throw the girl?'

(167) koʃ-ujor-mu-sun
run-PROG-Q?-2SING
'are you running?'

(173) kitab-ɨ oku-jadʒak-mɨ-jɨz
book-OBJ read-FUT-Q?-1PL
'will we read the book?'

This is the only real irregularity in Turkish interrogatives. When asking questions which expect specific information as an answer, a simple set of interrogative pronouns is used. These pronouns function (almost) exactly as nouns, and take all the regular nominal morphology. The interrogative 'who', for instance, is simply the morpheme *kim*. Like regular nouns, it takes no morphology in the subject position, object marking in the object position, and presumably standard morphology in other nominal positions as well. (While we do not have data on the 'who' morpheme for any other thematic roles in the sentence, we do have such data for our other interrogatives, and can assume that *kim* functions analogously). The noun being 'questioned' is simply replaced with the interrogative pronoun, and no movement is required. There is, however, one instance in the data of the direct object being fronted while the interrogative is in the subject position. Compare (180) and (182) below:

(180) kim ajɨ-nɨn jɛmɛʜ-in-i ʧal-dɨ
who bear-POSSR food-POSSD-OBJ steal-PAST
'who stole the bear's food?'

(182) ajɨ-nɨn jɛmɛʜ-in-i kim ʧal-dɨ
bear-POSSR food-POSSD-OBJ who steal-PAST
'who stole the bear's food?'

Next, the interrogative *hangi* is roughly equivalent to the English 'which?', and it replaces a DET in the sentence, as in (201):

(201) hangi kɛdi-ji øldyr-dy
which cat-OBJ kill-PAST
'which cat did he kill?'

The 'which' interrogative does not appear to take any special morphology, merely preceding the noun it modifies. However we cannot be entirely sure as the data is scarce.

Finally, Turkish has what I will call the general interrogative *nɛ*, because this particle can be used to form multiple types of interrogatives, and would thus be the most common. In my transcriptions it will be noted as WH, and its best literal translation is simply 'what?'. Thus there is a set of three interrogatives – one for questioning animates, one for inanimates, and one for questioning specific information about nouns (i.e. for eliciting determiners).

nɛ's most basic use is shown in (179) and (186):

(179) nɛ ʧal-dɨ
WH steal-PAST
'what did he steal?'

(186) ayʃɛ nɛ(ji) øldyr-dy
Ayşe WH-(OBJ) kill-PAST
'what did Ayşe kill?'

As with most direct objects, the use of the object marker is optional and related to specificity (i.e. one includes it if they are discussing a specific referent, rather than a general entity). This optionality is shown in (186), and also shows how the *nɛ* morpheme takes regular noun morphology. The WH particle can also be used in complex constructions with locatives, forming the interrogative 'where?', as in examples (188), (191), and (193).

(188) kitap nɛrɛ-dɛ
book WH-AT
'where is the book?'

(191) køpɛk oʜlan-ɨn nɛrɛ-sin-i ɨsɨr-mɨʃ
dog boy-POSSR WH-POSSD-OBJ bite-HEAR
'where [on his body] did the dog bite the boy?'

(193) nɛrɛ-dɛn gɛl-di-n
WH-FROM come-PAST-2SING
'where did you (just) come from?'

Notice first that there seems to be some sort of internal morphology with the *nɛrɛ* morpheme, but that it is still clearly related to the general interrogative. These three examples show the flexibility of this interrogative in locative constructions, and rather ingeniously. The literal translation of (191), for instance, would be 'the dog bit the

boy's what?' but since this possessive construction is the one used for locatives, it is best interpreted as simply 'where?'.

Finally, the interrogative morpheme can be combined phrasally with *zaman* 'time' to mean 'when', as in (176):

(176)	nɛ	zaman	kitab-ɨ	oku-du-n
	WH	time	book-OBJ	read-PAST-2SING
	'when did you read the book?'			

It is worth noting here that the interrogative in (176) fits nicely into the AP slot in the VP as well.

5.5.2 Possession

Possessive constructions form form a very important role in Turkish, as their uses are many. Possession is marked on both the possessor (POSSR) and the thing possessed (POSSD), the former indicated by the morpheme {-(n)Vn}, the latter by the morpheme {-Vn}, with its allomorphs /-Vm/, /-Vn/, and /-sVn/ for first, second, and third person. The POSSR always immediately precedes the POSSD. At times, this means that the marking for both the POSSR and POSSD will be the same. Accordingly, the POSSR always precedes the POSSD. Here is one example of a simple sentence in which the agent is possessed by another entity:

(66)	ayʃɛ-nin	kɛdi-si	bir	mantar	jɛ-miʃ
	Ayşe-POSSR	cat-POSSD	QU	mushroom	eat-HEAR
	'Ayşe's cat ate a mushroom'				

Again, if we interpret the POSSR noun as a noun, then we must be willing to say that a noun can be modified by an NP.

In a moment we will turn to the myriad of uses of the possessive morphology. From the mere fact that the possessive morphemes can be used in so many circumstances, it seems obvious that these morphemes encompass more than just physical possession of one entity by another. Instead, they have developed a broader modifying purpose, and serve to relate a phrase head to its adjunct. There is indeed an underlying agreement process which covers more than just physical possession by animate entities.

Now we turn to the uses of this broader 'possessive' morphology.

5.5.3 Location

Turkish, in so far as we can tell from the data, has no such thing as a strict preposition. Instead, its prepositional concepts are indicated either through case marking on the noun (e.g. TO, FROM, AT, 'with') or by utilizing possession combined with a noun of the relevant semantics. Thus, the concept of 'under' is constructed as follows:

(58) aʜadʒ-ɨn alt-ɨn-da uju-muʃ
tree-POSSR bottom-POSSD-AT sleep-HEAR
'he slept under the tree'

Literally, we might translate this as 'he slept at the tree's bottom/underside'. This is the standard construction for indicating location in Turkish. However, locatives can be formed without possession as well, as in (43) and (79).

(43) køpɛʜ-i jɛr-ɛ koj-muʃ-lar
dog-OBJ ground-TO put-HEAR-3PL
'they put the dog on the ground'

(79) tabak-dan jɛmɛk al-mɨʃ
plate-FROM food take-HEAR
'she took food from the plate'

(43) translated literally would be 'they put the dog to the ground'. Neither of these utilize possession in order to express location. Therefore it is the locative cases (AT, TO, FROM) which primarily indicate location, but often a possessive construction is needed to further specify the location. Consider what (58) might look like without the possessive:

?(58) ʔaʜadʒ-da uju-muʃ
tree-AT sleep-HEAR
'he slept at the tree'

We could probably say with confidence that this is a grammatical sentence of Turkish, meaning something more like 'he slept in the area/vicinity of the tree', but it is vague as to the specific location. So in order to relay any specificity of location at all, possessives must be used.

5.5.4 Reflexives

Turkish reflexives revolve around the use of the reflexive term *kɛndi*, meaning roughly ‘own’ (as in ‘he broke his own foot’) or perhaps ‘self’. Semantically, it is easiest to describe the lexeme as a modifier attributing possession to the subject of the clause. Syntactically *kɛndi* appears bound to the subject of its containing clause. Using the possessive construction once again, one can combine the reflexive lexeme in a phrase to mean ‘oneself’, like in (202):

(202)	bɛn-im	kɛndi-m-i	sɛv-ijor-um
	I-POSSR	RFLX-1SING-OBJ	like-PROG-1SING
	‘I like myself’		

And this does very literally means ‘I like my self’, although it is difficult to express such a translation precisely. Because the POSSD morphology indicates the person of the POSSR, the POSSR noun is often simply dropped, leaving a construction like (204):

(204)	kɛndi-sin-i	sɛv-di
	RFLX-POSSD-OBJ	like-PAST
	‘she liked herself’	

The reflexive can also be used in non-possessive constructions, taking regular nominal morphology as follows:

(208)	kɛndi	kɛndi-n-ɛ	ʃarkɨ	søjlɛ-di
	RFLX	RFLX-POSSD-TO	song	sing-PAST
	‘he sang by himself’ (i.e. of his own initiative)			

(209)	(kɛndi)	kɛndi-sin-ɛ	ʃarkɨ	søjlɛ-di
	RFLX	RFLS-POSSD-TO	song	sing-PAST
	‘he sang a song to himself/(his own self)’			

(210)	kɛndi-sin-ɛ	hɛdi-jɛ	vɛr-di
	RFLX-POSSD-TO	gift-OBJ	give-PAST
	‘he gave a gift to himself’		

Complex reflexives like (208) are difficult to parse apart. The possessor is an implied third person entity (‘him’), and the second reflexive seems to modify the first, meaning something analogous to ‘he sang to his own self’. Generally, it is interesting that a reflexive can in a sense modify another reflexive.

5.5.5 Relative Clauses

Another use of possessive constructions is relative clauses. In RCs where the relativized noun is the direct object of the RC, the subject of the RC 'possesses' the verb, as in (234)

(234)	oʜlan-ɨn	sɛv-di-ʜi	køpɛʜ-i	kadɨn	ʧal-ɨjor
	boy-POSSR	like-PAST-POSSD	dog-OBJ	woman	steal-PROG
	'the woman stole the dog which the boy likes'				

This also holds true when the relativized noun is part of a locative construction because, as just discussed, locatives are really just object NPs with a possessive to specify the exact type of location. So (259) and (255) below in fact follows the same construction as (234), only with an added possessive modifier attached to the object.

(259)	bɛn-im	kɛs-di-ʜi-m	aʜaʧ-ɨn	alt-ɨn-da	tavʃan	uju-jor-du
	I-POSSR	cut-PAST-POSSD	tree-POSSR	bottom-POSSD-AT	rabbit	sleep-PROG-PAST
	'the rabbit was sleeping under the tree that I cut down'					

(255)	kirli	bulaʃɨk-lar-ɨ	koj-du-ʜun	masa	yst-yn-y	tɛmizlɛ-di-m
	dirty	dish-PL-OBJ	put-PAST-POSSD-2S	table	top-POSSD-TO	clean-PAST-1SING
	'I cleaned the table that you put the dirty dishes on'					

As mentioned previously, it is interesting to note that 'table' splits the relative clause in (255), whereas typically the entire relative clause would precede the noun it modifies. So while it is clear from the morphology that *masa* belongs to the main clause, 'prepositional' nouns seem to need to stay embedded within the RC. It is possible that, in most Turkish sentences, the noun of the main clause undergoes raising from its position in the RC; but because *masa* here is so far down the accessibility hierarchy, its movement may be blocked. But without a detailed generative analysis, this is of course merely speculative.

In RCs where the relativized noun is the subject of the RC, the syntax and morphology is somewhat different. Again, the RC immediately precedes the noun it modifies, and consists of RC → NP VP. But instead of a possessive construction, the verb appears to take a relative marker that permits no other morphology on the verb. And while this morpheme ({-Vn}) looks strikingly similar to the possessive, we know it is in fact a specialized relative marker because of how it changes the verb morphology and because there is no element within the sentence marked as the possessor of this verb. (251) is an example:

(251) mavi ɛlbisɛ-ji gij-ɛn kɨz
blue dress-OBJ wear-REL girl
'the girl who is wearing the blue dress'

It makes sense that when the relativized noun is the subject of the RC, it not be marked by the possessive, because such a phrase would be ambiguous. While the verb would still retain its POSSD marker, its subject could not take the POSSR marker, because it must instead mark for the thematic role of the main clause. The function-specific relative marker on the verb does a good job avoiding this ambiguity. Still, it should be of no surprise that even this relative marker is very similar to the marker for POSSR.

5.5.6 Complex Sentences

The possessive is used in one final construction (although it would be surprising if other uses of the possessive did not exist), and that is with complement clauses. They are formed in a manner similar to relative clauses, where the verb in the lower clause takes the POSSD morpheme, and the subject of the lower clause is the POSSR. So in (250), the subject of 'want' is the implied POSSR 'I'.

(250) ananas-lar-ɨ istɛ-di-ʜim-i bil-ijor-lar
pineapple-PL-OBJ want-PAST-POSSD-OBJ know-PROG-3PL
'they know I want pineapples'

Or more literally, 'they know (of) my wanting pineapples'. It is also extremely interesting to note that the complement verb is not marked by the object morpheme as one might expect, but rather by the TO case. We know this because it takes the vowel harmony for the TO morpheme, and not the object marker, as evidenced in (261):

(261) o iki oʜlan-lar-ɨn okula git-di-ʜin-ɛ inan-ɨjor-um
DET two boy-PL-POSSR school go-PAST-POSSD-TO believe-PROG-1SING
'I believe those two boys went to school'

In contrast to this, some verbs are marked with the FROM morpheme instead. Whether the complement verb receives a TO, FROM, or accusative case marker seems to be information coded into the verb of the main clause. So, just like in Latin where prepositions code for a particular case, in Turkish main verbs also code for a specific case in any complement phrase. Listed below is each elicited verb and its respective case.

kork 'fear'	FROM
yzyl 'regret'	FROM
onla 'realize'	TO
ist 'wish/want'	TO
inan 'believe'	TO
bili 'know'	OBJ

Table 7

Based on this limited data, the difference between verbs which encode TO and verbs which encode FROM appears entirely semantic. Verbs with positive meaning take the TO case, while verbs with negative meaning take the FROM case, and simple indicative verbs take object marking. It is not uncommon to see semantic differences like these expressed in spatial relations. In English, for example, phrasal verbs which are positive utilize *up* (e.g. 'cheer up', 'things are looking up'), while those with negative meanings utilize *down* (e.g. 'that's a downer', 'don't let life get you down').

While marking the complementized verb by using possession and locative case marking is the preferred method of forming complex sentences, there is also another construction, more similar to English. In this type of construction, the morpheme *-ki* is suffixed to the main verb, and the verb is then immediately followed by the complement sentence, which utilizes standard Turkish syntax. (260) is such an example (cf. (261) above to see the differences between the two types of complements).

(260)	inan-ɨjor-um-ki	o	iki	oʜlan-lar	okula	git-di-lɛr
	believe-PROG-1SING-CP	DET	two	boy-PL	school	go-PAST-3PL
	'I believe that those two boys went to school'					

Interestingly, the *ki* morpheme does not follow the pattern for vowel harmony, so it isn't likely part of the nominal morphology, perhaps allowing it to serve as a standalone complementizer. This morpheme appears to be a borrowing from Persian, hence it's disfavored status in Nihan's elicitations (she regularly said that using the *ki* morpheme was the less natural of the two ways of forming complentizer phrases).

As mentioned, the CP always follows the main verb, but this is fairly unusual for Turkish, as the complement (at least in all the unmarked forms of the constructions we've seen) always precedes its head. However, if the CP were to come before the main verb, it would no doubt be easily confused with the arguments of the main clause, a likely explanation for this irregularity.

Both relative clauses and complement clauses straddle the hazy line between noun and verb. Complement clauses seem to more nominal than verbal (in that they usually translate best as gerunds and follow typical noun morphology), whereas relative clauses have some exceptional morphology, but very closely mimick a normal Turkish sentence. Relative clauses in Turkish are, in some sense, sentences that have become nominalized, but remain sentences in every other respect. Whatever the case, there is indeed a continuum between nominals and verbals in language, and relatives and complement clauses in Turkish fall somewhere in the middle.

§6. Independent Project

6.1 Introduction

There is a great deal of typological variation across languages as to how spatial relations between entities (e.g. 'in', 'on', 'above', 'beside') are conveyed. Some languages code spatial concepts into the verb stem (e.g. 'come', 'go'), or by verb affixation (for example, in Swahili *peleka* 'send' → *pelekea* 'send to'). Others, like English, use adpositions, and still others mark spatial relations on the noun. Most languages use some combination thereof. Turkish is interesting in that it uses a somewhat complex construction, typically formed using a spatial noun (e.g. 'top', 'side') and possessive morphology, as below:

(6) para dʒyzdan-ɨn itʃ-in-dɛ
money wallet-POSSR inside-POSSD-AT
'the money is inside the wallet' (lit. 'the money is at the wallet's inside')

At other times, simple locative cases can be used:

(14) iʃarɛt duvar-a
sign wall-TO
'the sign is on the wall'

Languages also differ greatly in what *types* of spatial relations they mark for. Choi & Bowerman (2001) have studied the ways in which languages divide up spatial semantic fields. They discovered, for example, that there is a universal spatial semantic continuum of situations from the most prototypical 'on' situation, to the most prototypical 'in' situation. Figure 2, taken from Choi & Bowerman, shows some of the situations on that continuum.

Figure 2 can be read as a gradient from the most 'on-like' situations on the left to the most 'in-like' situations on the right. In English, situations (a)-(e) fall under our

'on' concept; but in Dutch, for example, 'on' is divided into two separate types of spatial relations: 'op' for situations (a) and (b) and 'aan' for (c)-(e). This continuum of spatial situations is also universal in that it is implicationally scaled. That is, languages differ on where they make the dividing lines between different types of spatial terms (e.g. which situations count as 'on', and which count as 'in' or 'beside'?), and how many spatial terms they use, but the terms will always cover adjacent segments of the continuum. Different languages focus on different features of a situation for deciding which spatial term to assign it. In Choi & Bowerman's own words, "languages focus on surprisingly different properties for calculating whether situations qualify as instances of the same or different semantic categories of space…languages agree on the overall typology of the semantic space to be partitioned, but differ dramatically in how they work out the boundaries between neighboring categories" (p. 480).

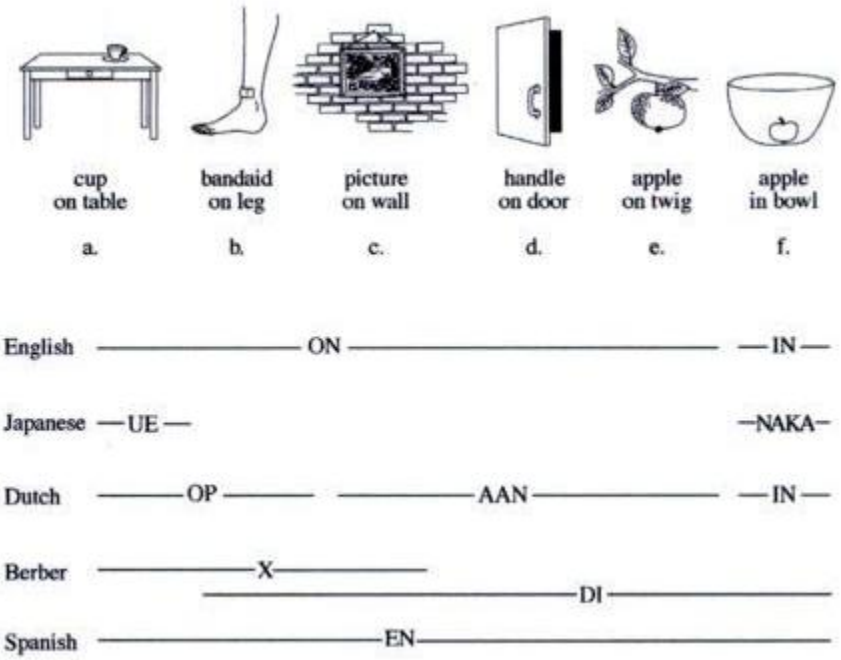


Figure 2

This project aims to examine the spatial semantics of Turkish, to understand which situations its spatial terms encompass, and what details of a situation are salient in distinguishing between spatial terms. I will focus on four types of relations specifically: TOP, BOTTOM, SIDE, and IN. For each of these I will also distinguish between situations involving contact and those without. A typical TOP without contact situation might be 'the book is above the table', whereas a SIDE with contact situation might be 'the painting is on the wall'.

6.2 Methodology

For this project I met with our informant Nihan one-on-one for two sessions of half an hour each, spaced about a week apart. In order to avoid cross-linguistic influence in a topic where modes of conceptualization are so important, I decided on a methodology that avoids translation as much as possible. For the first session, I brought with me a slew of objects and set them up in the immediate area, creating a field of interconnected spatial situations. Some objects and their placements were meant to elicit specific types of spatial relations (a box containing two metal balls for IN; a small sign for SIDE with contact; a Lego™ in the middle of a pile of paperclips, to see which spatial term it falls under; a wallet on the floor under the table). After setting up the field of objects, I then pointed out specific objects and asked Nihan to describe their location to me. This data was spontaneous and thus more natural, and even brought out some surprising features I will mention later. I then focused my questioning, and asked Nihan to describe the location of object x in relation to object y. While still avoiding direct translation, this allowed me to elicit data on specific spatial situations of interest.

The problem with the first session, oddly enough, was context – or rather, too much of it. Nihan would often resort to including additional details about the context in order to disambiguate the spatial relation, rather than utilize different spatial terms. In an effort to force the use of spatial terms (admittedly in opposition to more natural speech), the second session consisted eliciting sentences about hand-drawn pictures of objects rather than the objects themselves. This cut out a majority of the context that Nihan would typically have to rely on.

6.3 Data Analysis

The 48 elicited sentences for this project are given under Data at the end of this section. Most include a description of the spatial situation involved (rather than trying to recreate my not-so-professional artwork).

These data confirms the data from Choi & Bowerman. In situations similar to those in Figure 2, we find that the implicational scale holds. Turkish uses the spatial noun *yst* for situations (a) and (b), *itf* for (f), and locative cases (specifically the AT morpheme, {-dV}) for all others. This set of semantic divisions exactly mirrors those of Japanese, as shown in Figure 2.[6] Some examples of the relevant situations are shown here:

[6] It is interesting to note that the informant did once mention that Turkish is very similar to Japanese structurally, and that she can often understand segments of the language.

(8) dʒyzdan kutu-nun yst-yn-dɛ
wallet box-POSSR top-POSSD-AT
'the wallet is on the box'
Situation: TOP with contact

(30) saat bilɛʜ-in yst-yn-dɛ
watch wrist-POSSR top-POSSD-AT
'the watch is on the wrist'
Situation: SURROUND with contact

(14) iʃarɛt duvar-a
sign wall-TO
'the sign is on the wall'
Situation: SIDE with contact

(6) para dʒyzdan-ɨn itʃ-in-dɛ
money wallet-POSSR inside-POSSD-AT
'the money is inside the wallet'

It seems that Turkish makes a distinction between all the types of spatial situations studied here, except IN with contact and IN without contact – both simply use *itʃ*. Table 8 summarizes the uses of the various spatial terms. The term *orta* might also represent some form of containment. An example is given below.

Table 8

	Contact	No Contact
TOP	yst 'top'	jukarɨ 'upper part'
BOTTOM	alt 'bottom'	aʃaʜɨ 'bottom part'
SIDE	TO/AT	jan 'side' / karʃɨ 'across'
IN	itʃ 'inside'	
	orta 'middle'	

(22) lɛgo ataʃ-lar-ɨn orta-sɨn-da
Lego™ paperclip-PL-POSSR middle-POSSD-AT
'the Lego™ is in the middle of the paperclips'

The data gets somewhat messy with SIDE without contact, so that no single spatial term seems to be the favored. It is also interesting to note that *jukarɨ* and *aʃaʜɨ* didn't appear in any of the elicited sentences, but were given to me by Nihan instead. These were both very unnatural for her, and she would go to great lengths to avoid them. Consider (31)-(34) below:

(31) top kutu-nun yst-yn-dɛ
ball box-POSSR top-POSSD-AT
'the ball is on the box'

(32) top kutu-nun yst-yn-dɛ ama kutu-ja dokunujor
ball box-POSSR top-POSSD-AT but box-OBJ touching
'the ball is above the box but touching the box'

(33) (haljum dolu) balon kutu-nun yst-yn-dɛ utʃujor
helium fill balloon box-POSSR top-POSSD-AT flying
'the helium-filled balloon is hovering over the box'

(34) balon kutu-nun yst-yn-dɛ ama hava-da
balloon box-POSSR top-POSSD-AT but air-AT
'the balloon is on the box but in the air'

It seems from the data that *yst* is most naturally used for TOP with contact, but (32) and its translation of *yst* as 'above' suggests it might be most naturally construed as TOP without contact, and in fact Nihan seemed to lean towards this interpretation. She explained that, since objects don't normally hover, one generally assumes that *yst* implies contact. This leads one to expect that, with objects which normally *do* hover, *yst* alone (without additional specifications of context) more naturally implies TOP without contact. Yet when eliciting sentences about such a hovering object (a balloon), Nihan still depended on context to make the spatial situation clear (cf. (34)). Even in situations of TOP with contact, if pressed to indicate whether the situation is contact or no-contact the phrase still requires the addendum of a contextual phrase (e.g. 'touching the box' in (32) above). Nihan also found herself entirely unable to form a sentence using *jukarɨ* 'above', despite having given me this word as the most literal translation of above. These data suggest that *yst* is ambiguous for contact, and that when the indication of contact is relevant to discourse, it will be specified using, not more specific spatial terms, but with an additional contextual phrase.

It became increasingly clear with each example that the terms *yst* and *alt* were in fact ambiguous for contact. Nihan herself told me several times that sentences like (31) were ambiguous in this way. If we look again at Table 8, we notice that, while technically the terms do allow for an inherent contact/no-contact distinction, this doesn't seem apparent from the data. If *jukarɨ* and *aʃaʜɨ* are in fact never used, and if SIDE without contact requires choosing between a variety of terms, then it seems clear that Turkish spatial terms are inherently ambiguous for contact. *Yst, itʃ,* and *alt* require

the use of context to disambiguate, and SIDE is expressed in a variety of ways that make the context clear.

One final phenomenon of note was the appearance of the *ki* morpheme previously seen in complementizer clauses and once or twice with locatives. This was during the elicitation of the more spontaneous data. Examples (20) and (24) show this morpheme in use.

(20) ɛv-lar-dɛ-ki
house-PL-AT-?
'those in the house'

(24) ataʃ-ɨn jan-ɨn-da-ki lego jɛʃil-dir
paperclip-POSSR side-POSSD-AT-? Lego™ green-COP
'the Lego™ that is next to the clip is green'

The *ki* morpheme seems to have some sort of relativizing or modifying function, although clearly different from the function of usual relative clauses, because they typically involve entire phrases. So it seems best to suggest that this morpheme is a specifier of sorts, denoting additional details about the situation.

6.4 Data

1. onun itʃindɛ
   'inside it'

2. janjana
   'side by side'

3. jan
   'side'

4. ystystɛ
   'top on top', 'stacked'

5. ataʃlar janjana
   'the paperclips are side by side'
   Situation: IN, contact?

6. para dʒyzdanɨn itʃindɛ
   'the money is inside the wallet'
   Situation: IN, contact

7. kalam kalam kutusunun itʃindɛ
   'the pen is inside the pen-box'
   Situation: IN, no contact

8. dʒyzdan kutunun ystyndɛ
   'the wallet is on the box'
   Situation: TOP, contact

9. toplar kutunun itʃindɛ saklɨ
   'the balls are hidden inside the box'
   Situation: IN, contact

10. kalam kalam kutusunun itʃindɛ saklɨ
    'the pen is hidden inside the pen-box'
    Situation: IN, contact-ambiguous

11. toplar kutunun itʃindɛ
    'the balls are inside the box'
    Situation: IN, contact

12. iʃarɛt duvara asɨlɨ
    'the sign is hanging on the wall'
    Situation: SIDE, contact

13. iʃarɛt duvarda (asɨlmɨʃ)
    'the sign is (hanging) on the wall'
    Situation: SIDE, contact

14. iʃarɛt duvara – most natural
    'the sign is on the wall'
    Situation: SIDE, contact

15. kardɛʃim okulda
    'my sister is at school'
    Situation: general location

16. ʤyzdan masanɨn altɨnda
    'the wallet is under the table'
    Situation: BOTTOM, no contact

17. ʤyzdan masanɨn altɨna japɨʃmɨʃ
    'the wallet is stuck to the bottom of the table'
    Situation: BOTTOM, contact

18. ʤyzdan masanɨn ystyndɛ tutujorrum
    'I'm holding the wallet above the table'
    Situation: TOP, no contact

19. ystymyzdɛki tavanda asɨlɨ ɨʃɨklar
    'the lights that are hanging on the ceiling above us'
    Situation: TOP, no contact

20. ɛvlardɛki
    'those in the house'
    Situation: general location

21. lego ataʃlarɨn ystyndɛ
    'the lego is on the paperclips'
    Situation: TOP, contact

22. lego ataʃlarɨn ortasɨnda
    'the lego is in the middle of the paperclips'
    Situation: IN?, no contact

23. ʤyzdan vɛ kutu ystystɛ konulmuʃ
    'the box and the wallet are placed one on top of the other'
    Situation: TOP, contact
    - This is ambiguous as to which object is on top of which

24. ataʃɨn janɨndaki lego jɛʃildir
    'the lego that is next to the clip is green'
    Situation: SIDE, no contact

25. ʤyzdanɨn itʃindɛki para jɛʃil(dir)
    'the money that is in the wallet is green'
    Situation: IN, contact

26. bulaʃɨkdan aldɨнɨ jɛmɛk gyzɛldir
    'the food that she took from the plate is good'

27. duvarɨn ystyndɛki iʃarɛt kɨrmɨzɨdɨr
    'the sign that is on the wall is red'
    Situation: SIDE, contact

28. duvarɨn ystyndɛki oнlan kɨrmɨzɨdɨr
    'the boy who is on the wall is red'
    Situation: TOP, contact

29. duvarɨn ystyndɛ oturan oнlan kɨrmɨzɨdɨr
    'the boy who is sitting on the wall is red'

30. saat^{h} bilɛнin ystyndɛ
    'the watch is on the wrist'

Situation: SURROUND?, TOP?, contact

31. top kutunun ystyndɛ
'the ball is above the box'
Situation: TOP, no contact

32. top kutunun ystyndɛ ama kutuja dokunujor
'the ball is above the box but touching the box'
Situation: TOP, with contact

33. (haljum dolu) balon kutunun ystyndɛ utʃujor
'a (helium-filled) balloon is hovering [flying] over the box'
Situation: TOP, no contact

34. balon kutunun ystyndɛ ama havada
'the balloon is above the box but in the air'
Situation: TOP, no contact

35. balon kutunun ystyndɛ
'the balloon is on the box'
Situation: TOP, with contact

36. rɛsim duvara asɨlɨ
'the picture is hanging on the wall'
Situation: SIDE, with contact

36B. rɛsim duvardajdɨ
'the picture was on the wall'
(hanging implied)
Situation: SIDE, with contact

37. adam kɨzla ajnɨ odadajdɨ
'he was in the same room as her'

38. sandaljɛ duvara dajalijdɨ
'the chair was leaning against the wall'
Situation: SIDE, contact

39. sandaljɛ duvarɨn karʃɨsɨnda
'the chair is across from the wall'
Situation: SIDE?, no contact

39B. duvarɨn janɨnda
'the chair is next to the wall'
Situation: SIDE, no contact

40. top kutunun altɨnda
'the ball is beneath the box'
Situation: BOTTOM, no contact

41. top kutunun altɨnda ama kutuja dokunujor
'the ball is under the box but touching the box'
Situation: BOTTOM, with contact

42. top kutunun itʃindɛ
'the ball is inside the box'
Situation: IN, no contact

43. top su altɨnda
'the ball is under water'
Situation: IN, contact, immersion

44. top sujun itʃindɛ
'the ball is in the water'
Situation: IN, contact, immersion

45. top torbanun itʃindɛ
'the ball is inside the bag'
Situation: IN, no contact

46. top torbanun itʃindɛ
'the ball is inside the bag'
Situation: IN, contact

47. jukarɨdɨ
'above'
Situation: TOP, no contact

48. aʃaнɨda
    'under, beneath'
    Situation: TOP, no contact

Bibliography

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: universal and language-specific in the acquistiion of spatial semantic categories. In M. Bowerman, & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 475-511). Cambridge: Cambridge University Press.

Carnie, A. (2002). *Syntax: A Generative Introduction* (2nd Edition ed.). Malden, MA, United States of America: Blackwell Publishing.

Kornfilt, J. (1987). Turkish and the Turkic Languages. In B. Comrie (Ed.), *The World's Major Languages* (pp. 619-644). New York, NY: Oxford University Press.

Whaley, L. J. (1997). *Introduction to Typology: The Unity and Diversity of Language.* Thousand Oaks, California: SAGE Publications.

Wikipedia contributers. (2007, October 31). *Preposition and postposition*. Retrieved November 3, 2007, from Wikipedia, The Free Encyclopedia: http://en.wikipedia.org/w/index.php?title=Preposition_and_postposition&oldid=168317409