

Language specific peculiarities Document for

Turkish as Spoken in Turkey

1. Dialects

The seven dialect regions defined for Turkish are:

West	Marmara Region
South West	Aegean Region
South	Mediterranean Region
South East	South Eastern Anatolia Region
East	East Anatolia Region
North	Black Sea Region
Central	Central Anatolian Region

Modern Standard Turkish is the common standard dialect. This is the national standard and is based on the dialect spoken around İstanbul.

2. Deviation from native-speaker principle

No special deviation.

3. Special handling of spelling

No special deviation.

4. Description of character set used for orthographic transcription

Turkish uses the Basic Latin unicode range which is U+0041-U+007A plus the Latin Extended-A which is U+0100-U+017F. The Latin Extended-A characters included are the following:

Unicode	Character	Unicode	Character
U+011E	Š	U+0131	ı
U+011F	š	U+015E	Ş
U+0130	İ	U+015F	ş

Other non-standard characters include:

Unicode	Character	Unicode	Character
U+00C7	Ç	U+00E7	ç
U+00D6	Ö	U+00F6	ö
U+00DC	Ü	U+00FC	ü

Note that because of these additional characters, care has to be taken when approaching case in Turkish with the letters I, ı, İ, and i. The following mapping should be followed:

- Upper case “İ” with a dot (U+0130) maps to lower case “ı” with a dot (U+0069)
- Upper case “I” without a dot (U+0049) maps to lower case “i” without a dot (U+0131)

Upper Case		Lower Case	
Unicode	Character	Unicode	Character
U+0130	İ	U+0069	ı
U+0049	I	U+0131	i

5. Description of Romanization scheme

None.

6. Description of method for word boundary detection

Word boundaries in the orthography are determined by localization of white spaces (blank, tab, etc).

7. Table containing all phones in the stipulated notation

The phonemic transcription of words in this database is based on the official SAMPA symbol set for Turkish (which can be found at <http://www.phon.ucl.ac.uk/home/sampa/turkish.htm>) with the addition of the glottal stop /ʔ/ and the voiceless alveolar tap /r̥/. The total number of phones is 43. There are 25 consonants, 1 semi-vowel, and 16 vowels. The example transcriptions also contain the relevant suprasegmental symbols for stress, syllable and word boundary marking. Word boundaries are used in compound words like numbers and also in acronyms and words containing a hyphen.

IPA	SAMPA	Turkish Letter	Example Word	Example Transcription
CONSONANTS				
p	p	p	kapa	k a . " p a
b	b	b	baba	b a . " b a
t	t	t	tatlı	t a t . " 5 1
d	d	d	dede	d e . " d e
k	k	k	kale	k a . " l e
g	g	g	gaga	g a . " g a
f	f	f	fes	" f e s
w	w	v	ver	" w e r `
s	s	s	ses	" s e s

z	z	z	zam	“ z a m
ʃ	S	ş	şişe	S i . “ S e
ʒ	Z	j	jet	“ Z e t
tʃ	tS	ç	çam	“ tS a m
dʒ	dZ	c	cici	dZ i . “ dZ i
ɣ	G	ğ	sağır	s a . “ G 1 r`
c	c	k	kâr	“ c a r`
ɟ	gj	g	gavur	g j a . “ w u r`
h	h	h	hep	“ h e p
m	m	m	müjde	m y Z . “ d e
n	n	n	ne	“ n e
ŋ	N	n	Ankara	“ a N . k a . r a
l	l	l	elli	e l . “ l i
ʈ	ʂ	l	kul	“ k u ʂ
ʔ	ʔ	n/a	kur'a	k u r . “ ʔ a
r	r	r	ara	a . “ r a
ɾ	r`	r	var	“ w a r`
j	j	y	yaz	“ j a z
VOWELS				
a	a	a	ak	“ a k

e	e	e	keçi	ce. "tSi
i	i	i	kil	"ci l
o	o	o	oku	o. "ku
u	u	u	kutu	ku. "tu
y	y	ü	kül	"cyl
ø	2	ö	göl	"gj 2 l
u	1	ı	akıl	a. "k 1 5
a:	a:	a	dahi	da: . "hi
e:	e:	e	eğri	e: . "ri
i:	i:	i	ilmî	il. "mi:
o:	o:	o	doğru	do: . "ru
u:	u:	u	Suriye	su: . ri. "je
y:	y:	ü	üğrüm	y: . "ry m
ø:	2:	ö	öğretmen	2: . ret. "men
u:	1:	ı	ıdır	1: . "d 1 r`
SUPRASEGMENTALS				
'	"	primary stress	See above example transcriptions	
.	.	syllable break	See above example transcriptions	
	#	word boundary	e-posta	"e # p o s . " t a

8. Complete list of all rare phonemes

Rare phonemes are /y:/, /1:/, /ʔ/ and /G/. The glottal stop /ʔ/ could be considered a foreign phoneme, as it only occurs in words of Arabic origin and is fading out of use in Modern Standard Turkish. Historically, the letter “ğ” was realized as the voiced velar fricative (SAMPA /G/, IPA /ɣ/). In modern standard pronunciation it lengthens the preceding vowel, is deleted, or is realized as a glide. However, due to its continued use in some dialects, it has been included in this phoneset for the purposes of producing dialectal variants.

8.1.List of allophones

The allophones included in the phone set are:

- N (allophone of /n/ - before a velar consonant)
- r` (allophone of /r/ - word finally)
- c (allophone of /k/ - in the environment of front vowels)
- gj (allophone of /g/ - in the environment of front vowels)

Note: /c/ and /gj/ are phonemic in some words occurring adjacent to the vowel /a/. Usually this is indicated in the orthography with a caret 'â'. These words are usually not of Turkish origin.

9. Other Language Specific Items

9.1. Spelling alphabet

Orthography	Usual pronunciation	Orthography	Usual pronunciation
A, a	" a:	M, m	" m e:
B, b	" b e:	N, n	" n e:
C, c	" dZ e:	O, o	" o:
Ç, ç	" tS e:	Ö, ö	" 2:
D, d	" d e:	P, p	" p e:
E, e	" e:	R, r	" r e:
F, f	" f e:	S, s	" s e:
G, g	" gj e:	Ş, ş	" S e:
Ğ, ğ	j u . m u . " S a k . gj e	T, t	" t e:
H, h	" h e:	U, u	" u:
İ, i	" i:	Ü, ü	" y:
I, ı	" 1	V, v	" w e:
J, j	" Z e:	Y, y	" j e:
K, k	" c e:	Z, z	" z e:
L, l	" l e:		

9.2. Table of Digits

Digit	Turkish
0	sıfır
1	bir
2	iki
3	üç
4	dört
5	beş
6	altı
7	yedi
8	sekiz
9	dokuz
10	on
20	yirmi
30	otuz
40	kırk
50	elli
60	altmış
70	yetmiş
80	seksen
90	doksan

100	yüz
1000	bin
1000000	milyon

10. References

Comrie, Bernard. (1997). Turkish Phonology in Alan S. Kaye (ed) *The Phonologies of Asia and Africa: (including the Caucasus)*. Winona Lake, Indiana: Eisenbrauns.

Göksel, Aslı & Kerslake, Celia. (2005). *Turkish: A comprehensive grammar*. New York: Routledge.

Various online references.