



Language Manual

HQ and HD Turkish

Language Manual: HQ and HD Turkish

Published 18 March 2011

Copyright © 2008-2011 Acapela Group

All rights reserved

This document was produced by Acapela Group. We welcome and consider all comments and suggestions. Please, use the *Contact Us* link at our website:

<http://www.acapela-group.com>

Table of Contents

1. General	1
2. Letters in orthographic text	2
3. Punctuation characters	3
3.1. Comma, colon and semicolon	3
3.2. Quotation marks	3
3.3. Full stop	3
3.4. Question mark	3
3.5. Exclamation mark	3
3.6. Parentheses, brackets and braces	3
3.7. Apostrophe	3
4. Other non alphanumeric characters	4
4.1. Non-punctuation characters	4
4.2. The ² and ³ signs	5
4.3. Symbols whose pronunciation varies depending on the context	5
5. Number Processing	7
5.1. Full number pronunciation	7
5.2. Leading zero	8
5.3. Decimal numbers	8
5.4. Currency amounts	8
5.5. Ordinal numbers	9
5.6. Arithmetic operators	9
5.7. Mixed digits and letters	9
5.8. Time of day	10
5.9. Dates	10
5.10. Phone numbers	11
5.11. Roman Numerals	13
5.12. Numbers with suffixes	13
6. How to change the pronunciation	14
6.1. User lexicon	14
6.2. Alternative spelling	14
7. Turkish Phonetic Text	15
7.1. Consonants	15
7.2. Comments on the phonetic symbols for consonants	16
7.3. Vowels	16
7.4. Comments to the phonetic symbols for vowels	16
7.5. Glottal stop	16
7.6. Lexical stress	17
7.7. Inter-word processes	17
7.8. Pause	17
8. Abbreviations	18
9. Web-addresses and email	20

List of Tables

4.1. Non-punctuation characters	4
7.1. Symbols for the Turkish consonants	15
7.2. Vowels	16
8.1. Abbreviations	18

Chapter 1. General

This document discusses certain aspects of text-to-speech processing for the Turkish text-to-speech system, in particular the different types of input characters and text that are allowed.

This version of the document corresponds to the High Quality (HQ) voice İpek, and the High Density (HD) voices Baris and Ceren.

Please note that the *User's Guide*, mentioned several times in the manual, is called *Help* in some applications.

Note: This language manual is general and applies to all Acapela Group HQ and HD Turkish voices specified above. One or more of the voices may be included in a certain Acapela Group product.

Note: For efficiency reasons, the processing described in this document has a different behaviour in some Acapela Group products. Those products are:

- Acapela TTS for Windows Mobile
- Acapela TTS for Linux Embedded
- Acapela TTS for Symbian



For these products, the default processing of numbers, phone numbers, dates and times has been simplified for the low memory footprint (LF) voice formats. Developers have the possibility to change the default behaviour from *simplified* to *normal* preprocessing by setting corresponding parameters in the configuration file of the voice. Please see the documentation of these products for more information. In the following chapters, each simplification will be described by the indication *[not SP]* following the description of the standard behaviour. The *SP* in the indication stands for *Simplified Processing*.

Chapter 2. Letters in orthographic text

Characters from A-Z, a-z and [Ö ö Ç ç Ş ş İ ı Ğ ğ Ü ü İ î Â â] may constitute a word. Certain other characters are also considered as letters, notably those used as letters in other European languages, i.e. ñ, å, é. These letters are not pronounced as in their native languages though, they are pronounced as regular *n*, *a*, *e* when occurring in a word.

However, when one of these letters stands on its own, they are read with an indication of the diacritic (modification of the letter), for instance "ô" is read "*inceltme işaretli o*".

Characters outside of these ranges, i.e. numbers, punctuation characters and other non-alphanumeric characters, are not considered as letters.

Chapter 3. Punctuation characters

Punctuation marks appearing in a text affect both rhythm and intonation of a sentence. The following punctuation characters are permitted in the normal input text string: , ; “ ” . ? ! () [] { }

3.1. Comma, colon and semicolon

Comma ',', colon ':' and semicolon ';' cause a brief pause to occur in a sentence, accompanied by a small rising intonation pattern just prior to the character.

3.2. Quotation marks

Quotes "" appearing around a single word or a group of words cause a brief pause before and after the quoted text.

3.3. Full stop

A full stop '.' is a sentence terminal punctuation mark which causes a falling end-of-sentence intonation pattern and is accompanied by a somewhat longer pause. A full stop may also be used in the whole number part of the digit strings and for ordinal numbers (see chapter *Number processing*). It can also be used in abbreviations (see chapter *Abbreviations*).

3.4. Question mark

A question mark '?' ends a sentence and causes question-intonation, first rising and then falling.

3.5. Exclamation mark

The exclamation mark '!' is treated in a similar manner to the full stop, causing a falling intonation pattern followed by a pause.

3.6. Parentheses, brackets and braces

Parenthesis '()', brackets '[]', and braces '{} ' appearing around a single word or a group of words cause a brief pause before and after the bracketed text.

3.7. Apostrophe

Apostrophe "'" should be used to add suffixes to the numbers. In other cases, it does not make any difference to the pronunciation

Chapter 4. Other non alphanumeric characters

4.1. Non-punctuation characters

The characters listed below are processed as non-letter, non-punctuation characters. Some are pronounced at all times and others are only pronounced in certain contexts, which are described in the following sections of this chapter.

Table 4.1. Non-punctuation characters

Symbol	Reading
/	slaş
\	ters slaş
©	te lif hakkı
™	ticari marka
®	tahütlü marka
°	derece
^	inceltme işareti
¨	çift üst nokta
=	eşittir
>	büyüktür
<	küçüktür
-	tire
+	artı
*	yıldız
÷	bölü
×	çarpı
±	artı eksi
%	yüzde
‰	binde
~	tilde
#	diy ez
§	Bölüm işareti
\$	dolar
£	paunt
¥	yen
€	yuro
¢	sent
½	yarım
¼	çeyrek
¾	dörtte üç
²	kare
³	küp
µ	mikro
β	beta

Symbol	Reading
&	ve
	çubuk
@	et
–	alt çizgi
f	kancalı fe
†	kama işareti
‡	çift kama
₺	para birimi işareti
 	kesik çizgi
ª	dişi sıralama işareti
¶	paragraf işareti
º	erkek sıralama işareti
„	çengel
¹	üstü bir

4.2. The ² and ³ signs

The reading of expressions with ² and ³ is:

Expression	Reading
mm ²	milimetre kare
cm ²	santimetre kare
m ²	metre kare
km ²	kilometre kare
mm ³	milimetre küp
cm ³	santimetre küp
m ³	metre küp
km ³	kilometre küp

4.3. Symbols whose pronunciation varies depending on the context

4.3.1. Hyphen

A hyphen '-' is pronounced *eksi* in two cases:

1. if followed by a digit and no other digit is found in front of the hyphen, i.e. as in the pattern -X but not in X-Y or X-Z where X, Y, and Z are numbers.
2. if followed by a digit and an equals sign '=', i.e. as in the pattern X-Y=Z. Space is allowed between digits, hyphen and equals sign.

In between words, the hyphen is not pronounced.

Examples:

ara-bul, çizgi-adam

Hyphen is also ignored in some valid date and time formats. (see chapters *Dates* 5.9 *Phone numbers* 5.10).

In all the other cases, the hyphen is pronounced as *tire*.

Expression	Reading	
-3	eksi üç	
44-3	kırk dört tire üç	
44-3=41	kırk dört eksi üç eşittir kırk bir	
15-20 Ekim	on beş tire yirmi ekim	[not SP]

4.3.2. Asterisk

Asterisk '*' is pronounced *çarpı* if part of a mathematical expression containing an equals sign '='. In other cases it is pronounced *yıldız*.

Expression	Reading
2*3	iki yıldız üç
2*3=6	iki çarpı üç eşittir altı
*bc	yıldız b c
b*c=	b çarpı c eşittir

4.3.3. Equals sign

Equals sign '=' is pronounced as *eşittir*.

Expression	Reading
2*3=6	iki çarpı üç eşittir altı
cb=bc	c b eşittir b c

4.3.4. Slash

The slash sign '/' is read *bölü* if the input matches the patterns $X/Y=Z$. If the pattern is X/Y , the expression is read as a fraction. The expressions may contain blanks. In certain date formats, as a delimiter between the parts of the date, '/' is never pronounced. In other cases '/' is read '*slaş*'

Expression	Reading
2/3=0,67	iki bölü üç eşittir sıfır virgül altmış yedi
2/3	iki bölü üç
10/02/1999	on şubat bin dokuz yüz doksan dokuz
cb/bc	c b slaş b c
http://www.acapela-group.com	h t t p iki nokta slaş slaş w w w acapela tire group nokta kom

Chapter 5. Number Processing

Strings of digits that are sent to the text-to-speech converter are processed in several different ways, depending on the format of the string of digits and the immediately surrounding punctuation or non-numeric characters. To familiarise the user with the various types of formatted and non-formatted strings of digits that are recognised by the system, we provide below a brief description of the basic number processing along with examples. Number processing is subdivided into the following categories:

Full number pronunciation
Leading zero
Decimal numbers
Currency amounts
Ordinal numbers
Arithmetic operators
Mixed digits and letters
Time of day
Dates
Telephone numbers

5.1. Full number pronunciation

Full number pronunciation is given for the whole number part of the digit string.

Example

2425	full number
2.425	full number
24,25	24 is a full number, 25 is the decimal part

Numbers denoting thousands, millions and billions (numbers larger than 999) may be grouped using space or full stop (not comma). In order to achieve the right pronunciation the grouping must be done correctly.

The rules for grouping of numbers are the following:

- Numbers are grouped in groups of three starting at the end.
- The first group in a number may consist of one, two, or three digits.
- If a group, other than the first, does not contain exactly three digits, the sequence of digits is not interpreted as a full number.
- The highest number read is 999999999999 (twelve digits). Numbers higher than this are read as separate digits.

Number	Reading
2580	iki bin beş yüz seksen
2 580	"
2.580	"
25800	yirmi beş bin sekiz yüz
25 800	"
25.800	"
2580350	iki milyon beş yüz seksen bin üç yüz elli
2.580.350	iki milyon beş yüz seksen bin üç yüz elli

Number	Reading
2,580,350	iki virgöl beş yüz seksen (pause) üç yüz elli
1000000000	bir milyar
123 456 789 012	yüz yirmi üç milyar dört yüz elli altı milyon yedi yüz seksen dokuz bin on iki
1123456789012	bir bir iki üç dört beş altı yedi sekiz dokuz sıfır bir iki

5.2. Leading zero

Numbers that begin with 0 (zero) are read as a whole numbers. The initial zero's are pronounced too.

Number	Reading
09253	sıfır dokuz bin iki yüz elli üç
020	sıfır yirmi

5.3. Decimal numbers

The full number part of the decimal number (the part before comma or full stop) is read according to the rules in the section *Full number pronunciation*. If the decimals (the part after comma or full stop) are more than three, the decimal part is read as separate digits.

Number	Reading
16,234	on altı virgöl iki yüz otuz dört
3,1415	üç virgöl bir dört bir beş
1251,04	bin iki yüz elli bir virgöl sıfır dört
2,50	iki virgöl elli
2.51	iki nokta elli bir
3,141	üç virgöl yüz kırk bir
3.141	üç bin yüz kırk bir

5.4. Currency amounts

The following principles are followed for currency amounts:

- Numbers with zero or two decimals preceded or followed by the currency markers £, \$, ¥, €, YTL, TRY, TL are read as monetary amounts.
- Accepted decimal markers are comma ',' and full stop '.'.
- The decimal part (consisting of two digits) in monetary amounts is read as *yeni kuruş* for YTL and TRY, *kuruş* for TL, *peni* for pound, and *sent* for dollar and euro.
- If the decimal part is 00 it will not be read.

Expression	Reading
YTL 20.50	yirmi Yeni Türk Lirası 50 Yeni Kuruş
20,50 TRY	yirmi YTL 50 Ykr
\$15.00	on beş dolar
15.00£	on beş paunt
€15,00	on beş euro

Expression	Reading
€ 200,50	iki yüz euro elli sent
1.000.000 ¥	bir milyon yen
\$1.314,57	bin üç yüz on dört dolar elli yedi sent
\$ 1 milyon	bir milyon dolar

5.5. Ordinal numbers

Numbers are read as ordinals in the following cases:

- if they are followed by the ordinal suffix. Numbers followed by a full stop are not taken as ordinal numbers.
- If the number between two words is followed by a full stop. In other cases the full stop after the number will cause a pause.

Note: The roman numbers followed by a full stop are always taken as ordinal numbers, see section *Roman Numerals* 5.11.

Expression	Reading
1'inci	birinci
15. ev	on beş (pause) ev
Bu 15. ev	bu on beşinci ev

5.6. Arithmetic operators

Numbers together with arithmetical operators are read according to the examples below.

Expression	Reading
-12	eksi on iki
14-2=12	on dört eksi iki eşittir on iki
+24	artı yirmi dört
2*3=6	iki çarpı üç eşittir altı
4/2=2	dört bölü iki eşittir iki
2/3	iki bölü üç
%25	yüzde yirmi beş
%3,4	yüzde üç virgöl dört

5.7. Mixed digits and letters

If a letter appears within a sequence of digits, the groups of digits will be read as numbers according to the rules above. The letter marks the boundary between the numbers. The letter will also be read.

Expression	Reading
77B84Z3	yetmiş yedi b seksen dört z üç
0092B87-B	sıfır sıfır doksan iki b seksen yedi b
208YTL	iki yüz sekiz Yeni Türk Lirası
X2000	X iki bin

5.8. Time of day

Either colon ':', full stop '.', single quote ''' or double quote "" may be used to separate hours, minutes and seconds.

Possible patterns are:

- hh:mm* or *h:mm*
- hh:mm:ss* or *h:mm:ss*
- saat hh.mm* or *h.mm*

h = hour, *m* = minute, *s* = second.

For all of the patterns above, if the first digit of the hour is zero, it is omitted. Double zeros in the final part of the time expression are not pronounced either. Pattern c is recognised as a time expression if it is preceded by *saat* meaning hour.

Note: *00:00* is pronounced as *Gece yarısı*.

Expression	Reading
13:10	on üç on
01:13	bir on üç
1:30	bir otuz
13:15	on üç on beş
12:13:55	on iki on üç elli beş
12.11	on iki nokta on bir
Saat 12.11	saat on iki on bir
12:00:15	on iki sıfır sıfır on beş
12:00:00	on iki

5.9. Dates

There are 3 types of valid formats for dates:

- dd.mm.yyyy*, *dd/mm/yyyy*, and *dd-mm-yyyy*
- dd.mm.yy*, *dd/mm/yy*, and *dd-mm-yy*

yyyy is a four-digit number, *yy* is a two-digit number, *mm* is a month number between 1 and 12 and *dd* a day number between 1 and 31.

Hyphen, full stop, slash or space may be used as delimiters. In all formats, one or two digits may be used in the *mm* and *dd* part. Zeros may be used in front of numbers below 10. In these cases, zeros will not be pronounced.

Examples of valid formats and their readings:

Type 1:	Reading
03.02.2010 or 3.2.2010	üç Şubat iki bin on
03/02/2010 or 3/2/2010	üç Şubat iki bin on
03-02-2010 or 3-2-2010	üç Şubat iki bin on

Type 2:	Reading
03.02.10 or 3.2.10	üç Şubat iki bin on
03/02/10 or 3/2/10	üç Şubat iki bin on
03-02-10 or 3-2-10	üç Şubat iki bin on

Other possible formats include numbers combined with day and month names and their abbreviations.

Month names are only expanded if found in full date expressions, i.e. if preceded by a date and followed by a year.

Example	Reading
Pazartesi, 10 Şubat	Pazartesi, on Şubat
Cm, 10 Şubat 2003	Cuma, on Şubat iki bin üç
10 Şubat 2003	on Şubat iki bin üç
10 şub 2003	on Şubat iki bin üç

Day and month abbreviations are only expanded when appearing in correct date contexts. Month abbreviations are expanded if found in full date expressions, i.e. if preceded by a date and followed by a year. *cm* is expanded as a date abbreviation if it is not preceded by a digit, otherwise it is expanded as *santimetre*.

Valid abbreviations for months: *oca, şub, mar, nis, may, haz, tem, ağ, eyl, eki, kas, ara*.

Valid abbreviations for days: *pt, sl, çr, pr, cm, ct, pz*.

[not SP] Ranges of years are also supported.

Expression	Reading
1998-1999	bin dokuz yüz doksan sekiz doksan dokuz
1939-45	bin dokuz yüz otuz dokuz kırk beş

5.10. Phone numbers

In this section the patterns of digits that are recognized as phone numbers are described. In the pronunciation of phone numbers each group of digits is read as a full number with pauses between groups of numbers.

5.10.1. Ordinary phone numbers and mobile phone numbers

Sequences of digits in the following formats are treated as phone numbers:

- The regional code and mobile code consists of 3 digits, and is preceded by a zero.
- The actual phone number (henceforth the local number) consists of 7 digits with the only restriction that the first digit cannot be a zero.

Possible patterns of the national telephone numbers are shown below. The area or GSM codes can be separated by a space, a period, or a hyphen. The beginning zero can be in brackets. 'd' represents a digit.

- 0 ddd (area code or GSM code) ddd dd dd
- 0ddd (area code or GSM code) ddd dd dd
- (0)ddd (area code or GSM code) ddd dd dd

- ddd (area code or GSM code) ddd dd dd
- ddd dd dd

All zeros are pronounced. The delimiters (including space) result in a brief pause. Digits grouped together are read as cardinals.

Expression	Reading
216-458 96 32	iki yüz on altı - dört yüz elli sekiz - doksan altı - otuz iki
0216/458 96 32	sıfır - iki yüz on altı - dört yüz elli sekiz - doksan altı - otuz iki
0 216 458 96 32	sıfır - iki yüz on altı - dört yüz elli sekiz - doksan altı - otuz iki
0216 458 96 32	sıfır iki yüz on altı - dört yüz elli sekiz - doksan altı - otuz iki
(0)216 458 96 32	sıfır- iki yüz on altı - dört yüz elli - sekiz doksan altı - otuz iki
216 458 96 32	iki yüz on altı - dört yüz elli sekiz - doksan altı - otuz iki
458 96 32	dört yüz elli sekiz - doksan altı - otuz iki

5.10.2. Special phone numbers

There are also some numbers that don't need a regional code, for example:

The emergency number:	Reading
112	yüz on iki

5.10.3. International phone numbers

Phone numbers in other countries follow the pattern below (when calling from Turkey):

International Prefix + Country number + Regional number + Local number

<i>International Prefix:</i>	always 00
<i>Country number:</i>	1-3 digits
<i>Regional number:</i>	(0) + 1-3 digits
<i>Local number:</i>	5- digits

Example	Reading
0049(0)69 116 51	sıfır sıfır kırk dokuz – sıfır - altmış dokuz – yüz on altı – elli bir
0046 (0)8 799 86 19	sıfır sıfır kırk altı – sıfır - sekiz – yedi yüz doksan dokuz seksen altı – on dokuz

It is also common to replace the double zeros with a '+':

Example	Reading
+46 (0)8 799 86 19	artı kırk altı – sıfır - sekiz – yedi yüz doksan dokuz seksen altı – on dokuz

5.11. Roman Numerals

Certain letter combinations are interpreted as roman numerals, see the list below. Roman numerals are only recognized if they are written with capital letters. If they are followed by a dot, they are recognized as ordinals.

Example	Reading
III	üç
III.	üçüncü
IV	dört
VI	altı
VII	yedi
VIII	sekiz
IX	dokuz
XI	on bir
XII	on iki
XIII	on üç
XIV	on dört
XV	on beş
XVI	on altı
XVII	on yedi
XVIII	on sekiz
XIX	on dokuz

5.12. Numbers with suffixes

Suffixes are separated from the numbers by an apostrophe.

Example	Reading
12'sinde	on ikisinde
24'te	yirmi dördte
13'ü	on üçü

Chapter 6. How to change the pronunciation

6.1. User lexicon

Words that are not pronounced correctly by the text-to-speech converter can be entered in the user lexicon (see *User's guide*). When writing translations for entries in the user lexicon to change the way a word is pronounced, one method is to modify the spelling of the word (see section *Alternative spelling*) and another is to write a phonetic transcription of the word (see chapter *Turkish Phonetic Text*). Phonetic transcriptions can also be entered directly in the text, using the *PRN* or *PRX* tag (see *User's guide*).

6.2. Alternative spelling

Sometimes, the quickest way of changing the pronunciation of the word is to change the spelling of the word directly in the text. Changing a single letter, or adding a hyphen, can often make it sound better. This is specially useful when it comes to foreign words. Try to write the foreign words as they sound.

Correct spelling	Alternative spelling
Volkswagen	Volsvogen
Renault	Reno

Chapter 7. Turkish Phonetic Text

The Turkish text-to-speech system uses the Turkish subset of the SAMPA phonetic alphabet (*Speech Assessment Methods Phonetic Alphabet*) with some modifications. The symbols are written with a space between each phoneme.

Only the symbols listed here may be used in phonetic transcriptions. Symbols not listed here are not valid in phonetic transcriptions and will be ignored if included in the user lexicon or in a *PRN* or *PRX* tag.

7.1. Consonants

The table below lists the phonetic symbols used for the Turkish consonants along with example words and their transcriptions.

Table 7.1. Symbols for the Turkish consonants

Symbol	Word	Phonetic text	Comment
j	yat	j a1 t	
h	hasta	h a s t a1	
k	ak	a1 k	
c	kedi	c e d i1	
t	ütü	y t y1	
p	ip	i1 p	
b	bal	b a15	
d	dede	d e d e1	
g	karga	k a r g a1	
g'	genç	g' e1 n tS	
f	far	f a1 R	
s	ses	s e1 s	
tS	seçim	s e tS i1 m	
dZ	cam	dZ a1 m	
S	aşı	a S 11	
v	ver	v e1 r	
w	tavuk	t a w u1 k	Allophone of /v/
z	azık	a z 11 k	
Z	müjde	m y Z d e1	
l	lale	l a: l e1	
5	sal	s a1 5	
r	raf	r a1 f	Sonorant alveolar flap
R	kar	k a1R	Voiceless with friction allophone of /r/
m	dam	d a1 m	
n	anı	a n 11	
N	sanki	s a N c i1	
?	aslı	? a s 5 11	glotal stop

7.2. Comments on the phonetic symbols for consonants

- 'r before a pause is pronounced as /R/.
- /v/ is found only in word initial position , /w/ occurs in word internal and final position.
- /g/ and /c/ are found in the syllables with /e/ /i/ /y/ /2/ , as well as in some loan words with /a/ and /u/. For instance: *kağıt /c a 11 t/*
- /5/ never occurs in word initial position, it occurs in the syllables with /a/ /1/ /u/ /o/ in word internal and final positions. // occurs in word initial position with all vowels, in all syllables with /e/ /i/ /2/ /y/, and in some loan words with /a/ /1/ /u/ /o/. For instance : *gol /g o1 l/*
- Note that /1/ attached to a vowel is a stress mark and not part of a consonant or vowel. See section *Lexical Stress*.

7.3. Vowels

Table 7.2. Vowels

Symbol	Word	Phonetic text	Comment
i	döviz	d 21 v i z	
i:	iğne	i: n e1	
a	ortadan	o r t a d a1 n	
a:	ağ	a:1	
e	geçen	g e tS e1 n	
e:	eğri	/e: r i1	
u	ordusu	o r d u s u1	
u:	uğra	u: r a1	
o	Son	s o1 n	
o:	doğru	d o: r u1	
y	kül	k y1 l	
y:	AÜ	a: y:	
1	alınan	a 5 l n a1 n	
1:	sığ	s 1:1	
2	göre	g 2 r e1	
2:	öğretmen	2: r e t m e1 n	

7.4. Comments to the phonetic symbols for vowels

Long vowels are marked with colon ':'. For examples, see table above.

Note that '1' attached to a vowel is a stress mark and not part of a consonant or vowel. See section *Lexical Stress*

7.5. Glottal stop

A glottal stop, represented by the phonetic symbol 'ʔ' , is a small sound which is often used to separate two words when the first word ends and the second word starts with a vowel.

Example	Reading
beni ara (call me)	/b e n i 1 ? a r a 1/

7.6. Lexical stress

Lexical stress is used to indicate the level of prominence (or emphasis) of a syllable in a word. In words with more than one syllable, one of the syllables is more prominent than the others. This is referred to as word stress, or lexical stress. Words of one syllable also have word stress when spoken in isolation, although many may lose the stress in certain contexts. For the correct pronunciation of a word, it is important to include stress marks when writing phonetic transcriptions.

In phonetic transcriptions, lexical stress is indicated by the symbol '1' attached to the stressed vowel. Only vowels are stressed, i.e., a stress mark can only appear on a vowel, never on a consonant.

Turkish words have stress mostly on the last syllable. There are few exceptions of suffixes, place names and loan words.

7.7. Inter-word processes

Some changes in the phonetic transcriptions occur at the word boundaries:

If the last consonant of a word with at least two syllables is /p/, /tS/, /t/ and the next word starts with a vowel, the last consonant is voiced and pronounced as /b/, /dZ/, /d/ respectively. For example *şarap içiyorum* (*I am drinking wine*) would actually be pronounced as /S a r a b _ i tS i j o r u m/ instead of /S a r a 1 p _ i tS i 1 j o r u m/.

If the word preceding emphasis clitics, *de* /d e/ or *da* /d a/, ends with a voiceless consonant, then /d/ is pronounced as /t/. For instance, *kitap da* would be pronounced as /c i t a p _ t a/.

Glottal stop is inserted at word boundaries as explained in section *Glottal stop*.

7.8. Pause

An underscore /_/ in a phonetic transcription generates a small pause.

Chapter 8. Abbreviations

Some examples of the abbreviations that are recognized in the current version of the Turkish text-to-speech system are listed in the table below. The abbreviations marked with '*' are case-sensitive, the others are case-insensitive. I.e. the latter may be written with upper or lower case letters, the former have to be written as indicated in the table.

The abbreviations do not require a full stop in order to be recognized as abbreviations. The presence or absence of a full stop after the abbreviation should be observed. Some abbreviations representing units of measurement and measures of capacity are only expanded after digits. Abbreviations connected to telephony are only expanded in front of digits.

As previously mentioned, there are also abbreviations for the days of the week and the months (see chapter *Ordinal numbers* .)

Table 8.1. Abbreviations

Abbreviation	Reading
m	metre
l	litre
g	gram
w	vat
°c	santigrat derece
°f	fahrenayt derece
°k	kelvin derece
telefon no	telefon numarası
tel no	telefon numarası
tel	telefon
ev no	ev numarası
iş no	iş numarası
mm	milimetre
cm	santimetre (if preceded by a number)
dm	desimetre
hm	hektametre
km	kilometre
gr	gram
cg	santigram
dg	desigram
kg	kilogram
ml	mililitre
cl	santilitre
dl	desilitre
*kW	kilovat
km/saat	Kilometre saat
ft	fit
yd	yarda
n°	no

Abbreviation	Reading
*dB	desibel
*Kb	kilobayt
*Mb	megabayt
*Gb	gigabayt
bkz	bakınız
dk	dakika
yrd	yardım
yrb	yarbay
ltd	limited

Some acronyms and abbreviations are spelled in English if they are commonly used in that way. For instance *FBI* is pronounced as /e f b i: a1 j/.

Letter 'k' is pronounced as /k a/ or /k a:/, as in *SSK* /s e s e k a:1/ if the acronym is commonly used in that way, and if it is present in the dictionary. In all the other cases, it is pronounced as /c e/.

Some common acronyms that are written without diacritics will be recognised, if they are in the system.

Example	Reading
IMF	/i: m e: f e:1/
ODTU	/o d t y1/

Chapter 9. Web-addresses and email

Web-addresses and email-addresses are read as follows:

- *www* is spelled letter by letter.
- Full stops are read as *nokta*, hyphens as *tire*, underscore '_' as *alt çizgi*, slash '/' as *slaş*.
- *tr*, *uk*, *us* and all the other abbreviations for countries are spelled out letter by letter.
- The '@' is read *et*.
- Words/strings (including *org*, *com* and *edu*) are pronounced according to the normal rules of pronunciation in the system and in accordance with the lexicon.

String

`www.acapela-group.com`

`http://www.acapela-group.com`

`sales@acapela-group.com`

`support@acapela-group.com`

Reading

w w w nokta acapela tire group nokta com

h t t p iki nokta slaş slaş acapela tire group nokta com

sales et acapela tire group nokta com

support et acapela tire group nokta com