Language Manual

HQ Gothenburg Swedish

Language Manual: HQ Gothenburg Swedish

Published 20 September 2011
Copyright © 2011 Acapela Group.

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

List of Tables

Chapter 1. General

This document discusses certain aspects of text-to-speech processing for the Gothenburg Swedish 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 Kal.

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 Gothenburg Swedish 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-Ö* and *a-ö* may constitute a word. Certain other characters are also considered as letters, notably those used as letters in other European languages, i.e. *é, è, ô, ü*. These letters are not pronounced as in their native languages though, they are pronounced as regular *e, o, u* etc.

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 as a decimal marker in a number (see chapter *Number processing*) and 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.

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
/	snedstreck
+	plus
$	dollar
£	pund
€	euro
¥	yen
<	mindre än-tecken
>	större än-tecken
%	procent
^	cirkumfleks
|	lodtecken
~	tilde
@	snabel-a
²	(see below)
³	(see below)
*	(see below)
-	(see below)
=	(see below)

4.2. The ² and ³ signs

The reading of expressions with ² and ³ is:

Expression	Reading
mm²	kvadratmillimeter
cm²	kvadratcentimeter
m²	kvadratmeter
km²	kvadratkilometer
mm³	kubikmillimeter
cm³	kubikcentimeter
m³	kubikmeter
km³	kubikkilometer

4.3. Symbols whose pronunciation varies depending on the context

4.3.1. Hyphen

A hyphen '-' is pronounced *minus* 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.

If there is no equals sign, as in X-Y or X -Z, the hyphen is pronounced *streck*.

In certain date formats, in between days or years, the hyphen is pronounced *till*. In other cases the hyphen is never pronounced. Multiple occurrences of hyphen are pronounced *streck streck streck….*

Expression	Reading	
-3	minus tre	
44-3	fyrtiofyra streck tre	
44-3=41	fyrtiofyra minus tre är lika med fyrtioett	
44 - 3 = 41	fyrtiofyra minus tre är lika med fyrtioett	
15-20 oktober	femtonde till tjugonde oktober	[not SP]
6-10 nov	sjätte till tionde november	[not SP]
åren 1998-2004	åren nittonhundranittioåtta till tjugohundrafyra	[not SP]
2000-07-31	tjugohundra noll sju trettioett	
norsk-dansk	norskdansk	

4.3.2. Asterisk

Asterisk '*' is pronounced *gånger* only if the input matches the pattern X*Y=Z. In other cases it is pronounced *asterisk*.

Expression	Reading
2*3	två asterix tre
2*3=6	två gånger tre är lika med sex
*bc	asterisk b c

4.3.3. Equals sign

Equals sign '=' is pronounced *är lika med* if preceded or followed by a digit. In all other cases it is pronounced *likhetstecken*. Examples:

Expression	Reading
2*3=6	två gånger tre är lika med sex
cb=bc	c b likhetstecken b c

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.

• An exception is made for year pronunciation, which occurs in four-digit strings in the range between 1100 and 2099, see section 5.9.

Number	Reading
2580	tvåtusen femhundraåttio
2 580	"
2.580	"
25800	tjugofemtusen åttahundra
25 800	"
25.800	"

Number	Reading
2580350	tvåmiljoner femhundraåttiotusen trehundrafemti
2 580 350	"
2.580.350	"
1000000000	en miljard
23 456 789 012	tjugotremiljarder fyrahundrafemtiosexmiljoner sjuhundraåttioniotusen tolv
1234567890124	ett två tre fyra fem sex sju åtta nio noll ett två fyra

5.2. Leading zero

Numbers that begin with 0 (zero) are read as separate digits. If an initial zero is followed by exactly two digits (where the first of these is not a zero) the two last digits are read as a number. An exception to this is 070 (mobile telephone prefix) which is read *noll sju noll*.

Number	Reading
09253	noll nio två fem tre
020	noll tjugo

5.3. Decimal numbers

Comma or full stop may be used when writing decimal numbers (for currency amounts, colon may also be used, see section *Currency amounts*).

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. Note: A number containing full stop followed by exactly three digits is not read as a decimal number but as a full number, following the rules in the section *Full number pronunciation* .

Number	Reading
16,234	sexton komma tvåhundratrettiofyra
3,1415	tre komma ett fyra ett fem
1251,04	ettusentvåhundrafemtioett komma noll fyra
2,50	två komma femtio
2.50	två punkt femtio
3.141	tretusenetthundrafyrtioett

5.4. Currency amounts

The following principles are followed for currency amounts:

• Numbers with zero, one, or two decimals preceded or followed by the currency markers *kr, £, $, ¥* or *€* are read as currency amounts.

• Numbers with zero, one, or two decimals followed by the words *kronor, pund, dollar, yen or euro* are read as currency amounts.

• Accepted decimal markers are colon ':', comma ','and full stop.

• The sequence colon followed by hyphen ':-' is not read.

• No spaces are allowed in the number.

- The decimal part (consisting of one or two digits) in currency amounts is read as *och nn öre*, *och nn pence*, and *och nn cent* respectively.

- If the decimal part is *00* it will not be read.

Expression	Reading	
kr 20:50	tjugo kronor och femtio öre	
20:50 kronor	tjugo kronor och femtio öre	[not SP]
$15:00	femton dollar	
15:00£	femton pund	
15:00 euro	femton euro	[not SP]
€ 200:50	tvåhundra euro och femtio cent	
1.000.000 ¥	en miljon yen	
$1.314,57	ett tusen trehundrafjorton dollar och 57 cent	

5.5. Ordinal numbers

Numbers are read as ordinals in the following cases:

- The number is followed by a month name or one of the month name abbreviations and the number is smaller or equal to 31. The number may be preceded by a day or an abbreviation for a day.

- The number consists of a day interval followed by a month name/abbreviation.

- The number is part of the date format *dd/mm yyyy* and occurs in the *dd/mm* part. *dd/mm* must be a possible date and *yyyy* a year between 1100 and 2099. See also section *Dates* .

- The number is followed by colon '*:*' and an '*e*' (or an '*a*' after 1 and 2).

  Examples: *1:e, 1:a, 5:e*.

The valid abbreviations for months are: *jan, feb, febr, mar, apr, jun, jul, aug, sep, sept, okt, nov, dec*.

The valid abbreviations for days are: *mån, tis, ons, tor, tors, fre, lör, sön*.

The abbreviations above are only expanded to names of months and days when appearing in correct date contexts.

Expression	Reading	
15 januari	femtonde janauari	[not SP]
15 jan	femtonde januari	[not SP]
ons 15 jan	onsdag femtonde januari	[not SP]
15-16 januari	femtonde till sextonde januari	[not SP]
3/7 2007	tredje i sjunde 2007	

5.6. Arithmetic operators

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

Expression	Reading
-12	minus tolv

Expression	Reading
14-2	fjorton streck två
14-2=12	fjorton minus två är lika med tolv
+24	plus tjugofyra
2+3	två plus tre
2+3=5	två plus tre är lika med fem
2*3	två asterix tre
2*3=6	två gånger tre är lika med sex
2/3	två tredjedelar
25%	tjugofem procent
3,4%	tre komma fyra procent

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	sjuttiosju B åttiofyra Z tre
0092B87-B	noll noll nio två B åttiosju B
208Kr	tvåhundraåtta kronor

5.8. Time of day

Numbers denoting time are marked by the abbreviation *kl* (with or without a full stop) or the word *klockan* in front of the digits. Either colon ':' or full stop '.' may be used to separate hours, minutes and seconds.

Possible patterns are:

a. *kl hh:mm* or *h:mm*

b. *kl hh:mm:ss* or *h:mm:ss*

Full stop '.' may be used instead of colon and *klockan* instead of *kl* in both patterns.

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

Pattern a: if the *mm*-part is something other than *00*, an *och* will be inserted before this part. If the *mm*-part is equal to *00*, no *och* will be inserted.

Pattern b: an *och* will always be inserted before the *ss*-part, regardless of its value. No *och* will be inserted before the *mm* part in this pattern.

Expression	Reading
kl 10.15	klockan tio och femton
kl 9.30	klockan nio och trettio
kl 10:15:35	klockan tio femton och trettiofem
kl 9.30.27	klockan nio trettio och tjugosju
kl 8.00	klockan åtta noll noll

5.9. Years

[not SP] Numbers between 1100 and 2099 are always read as hundreds (year reading) with the exception of numbers containing decimals.

Expression	Reading
året 2008	året tjugohundraåtta
åren 1939-45	åren nittonhundratrettionio till fyrtiofem
åren 1998-2010	åren nittonhundranittioåtta till tjugohundratio
år 2000	år tjugohundra
X2000	X tvåtusen
år 2007	år tjugohundrasju
1088	ettusenåttioåtta
1900	nittonhundra
1988	nittonhundraåttioåtta
2000	tjugohundra
1988,0	ettusenniohundraåttioåtta komma noll
1988.32	ettusenniohundraåttioåtta punkt trettiotvå
september 2007	september tjugohundrasju
sep 2008	september tjugohundraåtta
13 sep 2019	trettonde september tjugohundranitton

5.10. Dates

There are three types of valid formats for dates:

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

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

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

yyyy is a four-digit number between 1100 and 2099, *yy* is a two digit number, *mm* is a month number between 1 and 12 and *dd* a day number between 1 and 31.

In type *1* and *2*, hyphen, full stop, and slash may be used as delimiters, in type *3*, only slash.

In all three formats, one or two digits may be used in the *mm* and *dd* part. Zeros may be used in front of numbers below 10.

Examples of valid formats and their readings:

Type 1:

2003-02-10 or 2003-2-10	tjugohundratre (noll)två tio
2003.02.10 or 2003.2.10	"
2003/02/10 or 2003/2/10	"

Type 2:

10-02-2003 or 10-2-2003	tio (noll)två tjugohundratre
10.02.2003 or 10.2.2003	"
10/02/2003 or 10/2/2003	"

Type 3:

10/02 2003 or 10/2 2003	tionde i andra tjugohundratre
10/2 –03 or 10/02 –03]	[not SP]

"

5.11. Phone numbers

In this chapter the patterns of digits that are recognised as phone numbers are described. In the pronunciation of phone numbers each group of digits is read as a full number (see also *Leading zero* section) with a pause between the regional code and the local number, and pauses between groups of numbers. [not SP] Hyphens and parentheses are not read.

5.11.1. Ordinary phone numbers

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

• The regional code consists of 2-4 digits, the first digit is always a zero.

• The actual phone number (henceforth the local number) consists of 5-8 digits with the only restriction that the first digit cannot be a zero.

• The regional code is followed by hyphen or slash followed by the local number. There should be no spaces around the hyphen/slash.

The information above is summarised in the following tables:

Number of digits in regional code	Examples of regional codes	Possible number of digits in local number
2	08	6, 7, 8
3	011, 031	5, 6, 7
4	0121	5, 6

Number of digits in local number	Grouping	Examples
5	xxx xx	220 28
6	xx xx xx	14 27 95
7	xxx xx xx	668 01 50
8	xxx xxx xx	500 248 92

Combination (regional code + local number)	Examples
2+6	08-33 12 31
2+7	08-702 16 80
2+8	08-500 248 92
3+5	011-405 31
3+6	011-14 27 95
3+7	031-612 81 30
4+5	0121-220 28
4+6	0200-230 230 or 0200-23 02 30

5.11.2. Special phone numbers

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

The emergency number:	112
Telia's Customer Service:	90 200
Number information	118 118

Such numbers are not recognised as phone numbers by the system and are therefore pronounced according to the general rules for pronouncing full numbers. Note that 118 118 and similar numbers are recognised as phone numbers if preceded by a regional code. Example: 08-320 100 is read *noll åtta trehundratjugo etthundra.*

5.11.3. Mobile phone numbers

Mobile numbers always consist of 10 digits: a prefix consisting of 3 digits plus 7 other digits (the subscriber's number). Traditionally the first of these 7 digits was the operator's number but this is not a general rule anymore. The operator's number may either be grouped with the prefix or with the subscriber's number.

The grouping possibilities are:

xxx-xxx xx xx
xxxx-xxx xxx
xxxx-xx xx xx

Example	Reading
010-432 15 12	noll tio (pause) fyrahundratrettiotvå (pause) femton (pause) tolv
073-245 35 04	noll sjuttitre (pause) tvåhundrafyrtiofem (pause) trettifem (pause) noll fyra
070-240 35 04	noll sju noll (pause) tvåhundrafyrtio (pause) trettiofem (pause) noll fyra
0702-45 35 04	noll sju noll två (pause) fyrtiofem (pause) trettiofem (pause) noll fyra

5.11.4. International phone numbers

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

International Prefix + Country number + Regional number + Local number.

International prefix:	always *00*
Country number:	1-3 digits
Regional number:	(0) + 1-3 digits
Local number:	5-8 digits

Examples:
0049(0)69 116 51
0046-(0)8-799 86 19
0046 (0)8 799 86 19

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

+46-(0)8-799 86 19

+46 (0)8 799 86 19

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 *Gothenburg Swedish Phonetic Text*). Phonetic translations can also be entered directly in the text, using the PRN-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.

Correct spelling	Alternative spelling
extern	extärn
juice	jos
mik	mick
pub	pubb
isjakt	is-jakt
kråkägg	kråk-ägg

This strategy can also be useful when it comes to foreign words. Try to write the foreign words as they sound in Swedish.

Correct spelling	Alternative spelling
light	lajt
James	Djäjms
date	dejt
sun	san
knock	nock
photo	foto
chunk	tjank
leave	liv

Note that it is also possible to write transcriptions including some English and German sounds, see section *Foreign sounds.*

Chapter 7. Gothenburg Swedish Phonetic Text

The Gothenburg Swedish text-to-speech system uses the Swedish subset of the SAMPA phonetic alphabet (Speech Assessment Methods Phonetic Alphabet) with one exception, */u/* is used instead of SAMPA's */u0/*.

Symbols not listed here are not valid in phonetic transcriptions and will be ignored if included in the user lexicon or in a PRN tag.

7.1. Consonants

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

Table 7.1. Gothenburg Swedish consonants

Symbol	Word	Phonetic text	Comment
b	bil	b i:4 l	
d	dal	d A:4 l	
rd	bord	b u:4 rd	retroflex d
f	fil	f i:4 l	
g	gås	g o:4 s	
h	hal	h A:4 l	
j	jag	j A:4 g	
k	skal	s k A:4 l	
k_h	kal	k_h A:4 l	aspirated k
l	lös	l 2:4 s	
rl	karl	k A:4 rl	retroflex l
m	mil	m i:4 l	
n	nål	n o:4 l	
rn	barn	b A:4 rn	retroflex n
N	ring	r I4 N	not syllable-initial
p	spill	s p I4 l	
p_h	pil	p_h i:4 l	aspirated p
r	ris	r i:4 s	
s	sil	s i:4 l	
rs	fors	f O4 rs	retroflex s
S	sjuk	S }:4 k	
t	stall	s t a4 l	
t_h	tal	t_h A:4 l	aspirated t
rt	hjort	j U4 rt	retroflex t
rt_h	årtull	o:3 rt_h u1 l	aspirated retroflex t
C	tjock	C O4 k	not syllable-final
v	vår	v o:4 r	

Note that */4/*, */3/* and */1/* are stress marks and not part of a consonant or vowel. Also */2/* may be a stress mark in the last syllable of a compound. See section *Stress marks* .

7.2. Comments on the phonetic symbols for consonants

7.2.1. Retroflexes

The pronunciation of the consonants '*t*', '*d*', '*s*', '*l*' and '*n*' is changed when they are orthographically preceded by an '*r*'. The tip of the tongue is bent backwards against the hard gum and the '*r*' is not pronounced as a separate sound. These sounds are called supradentals or retroflexes. Compare the pronunciation of *bord* and *bod*, *fort* and *fot*. In phonetic text these sounds are written as the combination of the '*r*' and the following dental consonant, with no space in between them: */rd, rl, rn, rs, rt/*. Note that not all dialects of Swedish have retroflexes.

7.2.2. Aspiration

In Gothenburg Swedish, the voiceless stop sounds '*P*', '*T*' and '*K*' are typically aspirated in certain positions of the word. That is, they are followed by a "puff of breath". Aspiration occurs in the following contexts:

• Initially in all words if followed by

   i. a vowel

   ii. one of the consonants '*J*', '*L*', '*N*', '*R*', '*V*' followed by a vowel.

• In all grave and acute non-compound words in front of a primary stressed vowel unless '*P*', '*T*', '*K*' are preceded by '*S*'. One of the consonants '*J*', '*L*', '*N*', '*R*', '*V*' is allowed between the aspirated stop and the vowel. (See section *Stress marks* for information on stress).

• Morpheme[1] initially in compounds followed by a vowel. One of the consonants '*J*', '*L*', '*N*', '*R*', '*V*' are allowed between the aspirated stop and the vowel. The previous morpheme may end with an '*S*'.

• Within a morpheme in a compound followed by

   i. a primary stressed vowel

   ii. one of the consonants '*J*', '*L*', '*N*', '*R*', '*V*' followed by a primary stressed vowel
   The aspiration does not occur if '*P*', '*T*', '*K*' are preceded by '*S*'.

Aspiration does not occur in the following contexts:

• After '*S*' within the same syllable.

• In front of /@/

• In grave non-compound words in front of vowel with secondary stress.

The synthesis cannot always correctly predict when a voiceless stop is to be aspirated. The user can indicate the aspiration of a voiceless stop by placing /_h/ after the stop symbol.

Note that prefixes like *be-* and *för-*, etc do not affect the aspiration.

Examples:

Aspiration:

pall	/p_h a4 l/
plats	/p_h l a4 t s/
katt	/k_h a4 t/

---

[1]A morpheme is the smallest meaningful unit in the grammar of a language. Examples: katt|er|na|s, trä|platta|n|s, skal|potatis|en|s

Aspiration:

prov	/p_h r U4 v/
tall	/t_h a4 l/
tvång	/t_h v O4 N/
potatis	/p_h U t_h A:4 t I s/
träplatta	/t_h r E:3 p_h l a1 t a2/
skalpotatis	/s k A:3 l p_h U t_h A:1 t l2 s/
tegeltak	/t_h e:3 g @ l t_h A:2 k/
hustak	/h }: 3 s t_h A:2 k /
betala	/b e t_h A:4 l a/

No aspiration:

spar	/s p A:4 r/
skatt	/s k a4 t/
stall	/s t a4 l/
hoppa	/h O3 p a/
hoppat	/h O3 p a t/
ljusstake	/j }:3 s s t A:1 k @2/

Note that the retroflex */rt/* may be aspirated.

Example: *parti /p_h a rt_h i:4/*

7.3. Vowels

The table below lists the phonetic symbols used for the Gothenburg Swedish vowels along with example words and their transcriptions.

Table 7.2. Gotheburg Swedish vowels

Symbol	Word	Phonetic text	Comment
a	hall	h a4 l	
A:	hal	h A:4 l	
e	vett	v e4 t	
e:	vet	v e:4 t	
@	pojken	p O3 j k @ n	schwa vowel allophone
I	vitt	v I4 t	
i:	vit	v i:4 t	
U	bott	b U4 t	
u:	sol	s u:4 l	
u	buss	b u4 s	
}:	hus	h }:4 s	
Y	bytt	b Y4 t	
y:	syl	s y:4 l	
O	håll	h O4 l	
o:	hål	h o:4 l	
E	rätt	r E4 t	

Symbol	Word	Phonetic text	Comment
E:	säl	s E:4 l	
{	herr	h {4 r	pre-r allophone of E
{:	här	h {:4 r	pre-r allophone of E:
2	föll	f 24 l	
2:	föl	f 2:4 l	
9	förr	f 94 r	pre-r allophone of 2
9:	för	f 9:4 r	pre-r allophone of 2:

Note that */4/*, */3/* and */1/* are stress marks and not part of a consonant or vowel. See section *Stress Marks* .

7.4. Comments to the phonetic symbols for vowels

7.4.1. Vowel length

Long vowels are marked with colon ':'. For examples, see table *Gothenburg Swedish vowels* .

7.4.2. Ä and Ö in front of R

Ä and *Ö* have special phonetic symbols when, orthographically, they immediately precede *R*. In this position *Ä* and *Ö* have a more open pronunciation. For *Ö*, */9/* and */9:/* are used, and for *Ä*, */{/* and */{:/* are used. See table *Gothenburg Swedish vowels* for examples.

7.5. Stress marks

Stress is used to indicate the level of prominence of a syllable in a word (word level stress) or of a word in a sentence (emphasis and reduction). In Gotheburg Swedish two words that are spelled identically may have different meanings depending on where in the word the stress is found. It is therefore very important to include stress marks when writing phonetic text. Note that transcriptions may be read with the correct stress even if no stress marks are included, but this happens randomly and is nothing that can be relied on.

7.5.1. Word level stress

Gotheburg Swedish is a tone language and has two word level accents: *acute* and *grave* accent.

• For *acute* accent the symbol */4/* is placed directly after (no space) the stressed vowel.

• For *grave* accent, the symbol */3/* is placed directly after the first stressed vowel, but the secondary stress should not be marked.

Compound words usually have grave accent and the symbol */3/* is used in the first part of the compound and 1 for secondary stress in the last. Gothenburg Swedish also have accent */2/* on the last syllable of the compound to indicate the small rise in intonation that is typical for the dialect. If accent *1* falls on the last syllable, this becomes a */2/* instead.

Accent	Word	Transcription	Comment
acute	anden	a4 n d @ n	the duck
grave	anden	a3 n d @ n	the spirit
acute	banan	b a n A:4 n	banana
grave	banan	b A:3 n a n	the track

Accent	Word	Transcription	Comment
compound	andpastej	a3 n d p_h a s t e2 j	duckpaté
compound	flodhästarna	f l u:3 d h E1 s t a rn a2	the hippopotamuses

Some more examples of words with *acute* accent:

hopp, bok, sten, vis, två
fågel, vatten, böcker, händer, kommer, springer
huset, borden, institutet, universitetet
större, högre, lägre
engelsk, matematisk
flanera, spendera
NOTE: måndag, blåbär, trädgård, verkstad

Some more examples of words with *grave* accent:

hoppa, hoppade, hoppat
tavla, tavlor, dike, diken, spänne, gamma
flicka, flickor, stenar
vishet, dårskap, vitaktig, tvåfaldig
ordbok, spelhåla

7.6. Foreign sounds

A few non-Swedish sounds, which sometimes occur in the pronunciation of foreign words and names, are permitted in the transcriptions. The sounds in question are exemplified in table below.

Table 7.3. Foreign sounds

Phonetic Symbol	Example	Phonetic text	Language of origin
aa	abend	aa4 b e n t	German
D	this	D I4 s	English
T	think	T I4 N k	English
tS	church	tS 2:4 tS	English
dZ	junk	dZ a4 N k	English
w	webmaster	w e4 b m A: s t @ r	English
z	zest	z e4 s t	English

7.7. Glottal stops

A glottal stop, represented by the phonetic symbol */?/* , is a small sound which is often used to separate two words when the second word starts with a stressed vowel. It is also useful when transcribing abbreviations. This sound can be inserted in a transcription in order to improve the pronunciation.

Example:

AEG   ? A: ? e: g e:4

7.8. Pause

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

Chapter 8. Abbreviations

In the current version of the Gothenburg Swedish text-to-speech system, the abbreviations in table Abbreviations are recognised in all contexts. These abbreviations are case-insensitive and require no full stop in order to be recognised as abbreviations.

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.

Expression	Reading
10 ml	tio milliliter
20 cl	tjugo centiliter
30 dl	trettio deciliter
40 mm	fyrtio millimeter (*mm* is normally read as *med mera*)
50 dm	femtio decimeter
tel 011-12 34 56	telefon noll elva (pause) tolv (pause) trettiofyra (pause) [not SP] femtiosex
tfn 08-987 654 32	telefon noll åtta (pause) niohundraåttiosju (pause) [not SP] sexhundrafemtiofyra (pause) trettiotvå
mob 070-245 35 04	mobil noll sju noll (pause) tvåhundrafyrtiofem (pause) [not SP] trettiofem (pause) noll fyra

Two abbreviations should be written with spaces or full stops in them: *t ex/t.ex.* (swe: *till exempel*, eng: *for example*) and *bl a/bl.a.* (swe: *bland annat*, eng: *among other things*).

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

Table 8.1. Abbreviations

Abbreviation	Reading
allm	allmän
amp	ampere
ang	angående
ankn	anknytning
anm	anmärkning
avd	avdelning
bilj	biljett
bitr	biträdande
bost	bostad
ca	cirka
cm	centimeter
co	company
div	diverse
dr	doktor
dvs	det vill säga
dyl	dylikt
eftr	efterträdare

Abbreviation	Reading
enl	enligt
etc	et cetera
ev	eventuellt
exkl	exklusive
exp	expedition
fr	från
hkr	hästkrafter
hr	herr
inkl	inklusive
inst	institution
jr	junior
kbm	kubikmeter
kg	kilo
kl	klockan
km	kilometer
kr	kronor
kvm	kvadratmeter
mg	milligram
milj	miljon
mkr	miljoner kronor
mm	med mera (except after numbers, see above)
mr	mister
mrs	missis
msk	matsked
nr	nummer
omkr	omkring
osv	och så vidare
pga	på grund av
prel	preliminär
proc	procent
resp	respektive
sekr	sekreterare
stud	studerande
tekn	teknisk
tsk	tesked

Chapter 9. Web-addresses and email

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

- *www* is read as three normal *v*'s spelled letter by letter.

- Full stops are read as *punkt*, hyphens as *streck*, underscore _ as *understreck*, slash */* as *snedstreck*.

- *se, uk, us* and all the other abbreviations for countries are spelled out letter by letter.

- The @ is read *snabel-a* .

- 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	Reading
www.google.se	v v v punkt google punkt s e
http://sv.wikipedia.org	h t t p kolon snedstreck snedstreck s v wikipedia punkt org
support@acapela-group.com	support snabel-a acapela streck group punkt com