



Language Manual

HQ Norwegian

Language Manual: HQ Norwegian

Published 17 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
4. Other non alphanumeric characters	4
4.1. Non-punctuation characters	4
4.2. The ² and ³ signs	4
4.3. Symbols whose pronunciation varies depending on the context	5
5. Number Processing	6
5.1. Full number pronunciation	6
5.2. Leading zero	7
5.3. Decimal numbers	7
5.4. Currency amounts	8
5.5. Ordinal numbers	8
5.6. Arithmetic operators	9
5.7. Mixed digits and letters	9
5.8. Time of day	9
5.9. Years	10
5.10. Dates	10
5.11. Phone numbers	11
6. How to change the pronunciation	12
6.1. User lexicon	12
6.2. Spelling incorrectly	12
7. Norwegian Phonetic Text	13
7.1. Stress marks	13
7.2. Symbols for the Norwegian consonants	14
7.3. Vowels	16
7.4. Glottal stops	16
7.5. Pause	17
8. Abbreviations	18
9. Web-addresses and email	19

List of Tables

4.1. Non-punctuation characters	4
7.1. Symbols for the Norwegian consonants	14
7.2. Symbols for the Norwegian vowels	16
8.1. Abbreviations	18

Chapter 1. General

This document discusses certain aspects of text-to-speech processing for the Norwegian 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) voices Kari, Olav and Bente.

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 Norwegian 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-Å and a-å may constitute a word. Certain other characters are also considered as letters, notably those used as letters in other European languages, e.g. é, è, ô, ü. These letters are not pronounced as in their native languages though, they are pronounced as regular e, o, u etc. 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 (for the letters mentioned above): *e med akutt aksent, e med grav aksent, o med sirkumfleks aksent, tysk y.*

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
/	skråstrek
+	pluss
\$	dollar
£	pund
€	euro
¥	yen
<	mindre enn
>	større enn
%	prosent
^	cirkumfleks
	vertikal strek
~	tilde
@	krøllalfa
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
mm ²	kvadratmillimeter
cm ²	kvadratcentimeter
m ²	kvadratmeter
km ²	kvadratkilometer
mm ³	kubikmillimeter
cm ³	kubikkcentimeter
m ³	kubikmeter
km ³	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 *bindestrek*.

In certain dateformats, it is pronounced *til*. If more than one hyphen appear together only one is processed.

Expression	Reading	
-3	minus tre	
44-3	førtifire bindestrek tre	[Not SP]
44-3=41	førtifire minus tre er lik førtien	[Not SP]
44 - 3 = 41	førtifire minus tre er lik førtien	
15.-20. oktober	femtende til tjuende oktober	[Not SP]
6.-10. nov.	sjette til tiende november	[Not SP]
årene 1998-2004	årene nittenhundreogtåtte til tusenogfire	[Not SP]
2000-07-31	totusen minus sju minus trettien	[Not SP]
data-maskin	datamaskin	

4.3.2. Asterisk

Asterisk '*' is pronounced *ganger* if enclosed by digits and followed by equals sign '='. In other cases it is pronounced *asterisk*.

Expression	Reading
2*3	to asterisk tre
2*3=6	to ganger tre er lik seks
*bc	asterisk b c

4.3.3. Equals sign

Equals sign '=' is pronounced *er lik* if followed by a digit. In all other cases it is pronounced *likhetstegn*.

Expression	Reading
2*3=6	to ganger tre er lik seks
cd=dc	c d likhetstegn d c

Chapter 5. Number Processing

The text-to-speech system applies the so-called new counting principles for Norwegian, which means that, for instance, 27 is read as *tjuesju* and not as *syvogtyve*.

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
Years
Phone 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
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.

Number	Reading
2585	totusen femhundreogåttifem
2.585	"
2 585	"

Number	Reading
25700	tjuefemtusen sjuhundre
25.700	"
25 700	"
2090350	tomillioner nittitusen trehundreogfemti
2.090.350	"
2 090 350	"
1000000001	en miljard og en
3456789012343	tre fire fem seks sju åtte ni null en to tre fire tre
34 567 890 123	trettifiremilliarder femhundreogsekstisjummiljoner åttehundreognittitusen ethundreogtjuetre

5.2. Leading zero

Numbers that begin with *0* (zero) followed by a maximum of three digits (the first of which is not *0*) are read *null* and then the number as it would be without the *0*. Numbers that begin with *00* (zero zero) followed by a maximum of two digits (the first of which is not *0*) are read *null null* and then the number as it would be without the *00*. Other digit strings beginning with *0* are read out digit-by-digit.

Number	Reading
0753	null sjuhundreogfemtitre
020	null tjue
0053	null null femtitre
00753	null null sju fem tre
07253	null sju to fem tre

5.3. Decimal numbers

Comma is used when writing decimal numbers.

The full number part of the decimal number (the part before comma) is read according to the rules in the section *Full number pronunciation* . If the decimals (the part after comma) 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	seksten komma tohundreogtrettifire
3,1415	tre komma en fire en fem
1251,04	ettusentohundreogfemtien komma null fire
2,50	to komma femti
2.50	to punktum femti
3.141	tretusenethundreogførtien

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 *kr*, £, \$, ¥ or € are read as currency amounts.
- Numbers with zero or two decimals followed by the words *kroner*, *dollar*, *yen* or *euro* are read as currency amounts.
- Comma is the only accepted decimal marker.
- The sequence comma followed by hyphen ',-' is read as *kroner*
- The decimal part (consisting of two digits) in currency amounts is read as *og nn øre*, *og nn pence*, and *og nn cent*.
- If the decimal part is *00* it will not be read.

Example	Reading
kr 20,50	tjue kroner og femti øre
kr 20,00	tjue kroner
kr 20,-	tjue kroner bindestrek
20,50 kroner	tjue kroner og femti øre
\$15,00	femten dollar
15,00£	femten pund
€ 200,50	tohundre euro og femti cent
1.000.000 ¥	en million yen
\$1.314,57	ettusen trehundreogfjorten dollar og femtisju cent

5.5. Ordinal numbers

Numbers are read as ordinals in the following cases:

- The number is followed by a full stop and 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 Dates.

Valid abbreviations for months: *jan*, *feb*, *febr*, *mar*, *apr*, *mai*, *jun*, *jul*, *aug*, *sep*, *sept*, *okt*, *nov*, and *des*.

Valid abbreviations for days: *man*, *tirs*, *ons*, *tor*, *tors*, *fre*, *lør* and *søn*.

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

Examples:

Expression	Reading	
3. januar	tredje januar	[Not SP]
3. jan.	tredje januar	[Not SP]
onsdag 3. jan.	onsdag tredje januar	[Not SP]
ons. 3. jan.	onsdag tredje januar	[Not SP]
15.-16. januar	femtende til sekstende januar	[Not SP]
3/7 2003	tredje i sjuende totusenogtre	

5.6. Arithmetic operators

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

Expression	Reading	
-12	minus tolv	
14-2	fjorten bindestrek to	[Not SP]
14-2=12	fjorten minus to er lik tolv	[Not SP]
+24	plus tjuefire	
2+3	to plus tre	
2+3=5	to plus tre er lik fem	
2*3	to asterisk tre	
2*3=6	to ganger tre er lik seks	
2/3	to tredjedeler	
6/2=3	seks dividert med to er lik tre	
2:3=6	to dividert med tre er lik seks	
25%	tjuefem prosent	
3,4%	tre komma fire prosent	

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	syttisju B åttifire Z tre
0092B87-B	null null nittito B åttisju B
208Kr.	tohundreogåtte kroner

5.8. Time of day

Numbers denoting time are marked by the abbreviation *kl* or the word *klokken* in front of the digits. Either colon or full stop may be used to separate hours, minutes and seconds.

Possible patterns are:

- kl hh.mm* or *h.mm*
- kl hh.mm.ss* or *h.mm.ss*

Colon may be used instead of full stop and *klokken* instead of *kl* in both patterns. *h* = hour, *m* = minute, *s* = second.

Expression	Reading
kl 9:00	klokken ni null null
klokken 9:30	klokken ni tretti
kl 13:00	klokken tretten null null
klokken 12:00	klokken tolv null null
kl 0:00	klokken null null
klokken 20.15.34	klokken tjue femten trettifire
kl 20.15.34	klokken tjue femten trettifire
20.15.34	tjue femten trettifire

5.9. Years

Numbers between 1100 and 2099 are always read as hundreds (“year reading”) with the exception of numbers containing decimals.

Expression	Reading	
året 1988	året nittenhundreogåttiåtte	[Not SP]
årene 1939-45	årene nittenhundreogtrettini til førtifem	[Not SP]
årene 1998-2010	årene nittenhundreognittiåtte til totusenogti	[Not SP]
år 2000	år totusen	
X2000	X totusen	
år 2004	år totusenogfire	
1088	ettusenogåttiåtte	
1900	nittenhundre	[Not SP]
1988	nittenhundreogåttiåtte	[Not SP]
2000	totusen	
1988,0	ettusenihundreogåttiåtte komma null	
1988.32	ettusenihundreogåttiåtte punktum trettito	
sep 2004	september totusenogfire	[Not SP]
13. sep 2004	trettende september totusenogfire	[Not SP]

5.10. Dates

There are four types of valid formats for dates:

	Format	Example
1.	<i>dd/mm/yy</i>	25/12/04
2.	<i>dd/mm/yyyy</i>	25/12/2004
3.	<i>dd-mm yyyy</i>	25-12 2004
4.	<i>dd/mm yyyy</i>	25/12 2004

All the above examples are read as *tjuefemte i tolvte totusen og fire*.

yyyy is a four-digit number 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 full stop, and slash may be used as delimiters, in type 3, only hyphen and in type 4 only slash.

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.

5.11. Phone numbers

In this section 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 pauses pause between the regional code and the local number, and pauses between groups of numbers.

5.11.1. Ordinary phone numbers

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

- A regular Norwegian phone number always has 8 digits, divided into groups of two or four.
- Norwegian mobile phone numbers may be divided 3-2-3
- International numbers are preceded by *00* or *+* plus country code

Examples:

22 33 44 55

2233 4455

960 23 123 [Not SP]

+47 22 33 44 55

00 47 22 33 44 55

5.11.2. Special phone numbers

The emergency numbers: 111, 112, 113

Service numbers: 1800 (4-digit numbers)

There are more numbers like this and they are pronounced according to the general rules for pronouncing full numbers.

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 transcriptions 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 *Spelling incorrectly*) and another is to write a phonetic transcription of the word (see chapter *Norwegian Phonetic Text*). Phonetic transcriptions can also be entered directly in the text, using the *PRN* tag (see *User's guide*).

6.2. Spelling incorrectly

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.

Chapter 7. Norwegian Phonetic Text

The Norwegian text-to-speech system uses the Norwegian subset of the SAMPA phonetic alphabet (*Speech Assessment Methods Phonetic Alphabet*) with a few exceptions. The symbols are written with a space between each phoneme. Earlier text-to-speech converters developed by Babel-Infovox use a different phonetic alphabet, called RULSYS. As some users may be accustomed to this alphabet we also provide here the corresponding RULSYS-symbols for every SAMPA-symbol.

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* tag.

7.1. Stress marks

Every Norwegian word of more than one syllable spoken in isolation has at least one (and normally only one) syllable that stands out as being more prominent than the others. This is referred to as *stress* (*word stress*, *lexical stress*). This stress also affects a word of one syllable pronounced in isolation. In an utterance of more than one word some words will retain the word stress, whereas some may lose theirs. For instance, so-called *function words* (pronouns, prepositions and other “grammatical” words) often lose their word stress in sentences. For instance, in the sentence *Jeg fant den på gaten* the words *jeg*, *den* and *på* will normally be unstressed. In some contexts even other words may lose their word stress. This can happen if another word receives a special emphatic stress, for instance, or in some regular word combinations, like *kjøre bil*, where *kjøre* is often pronounced without stress. In the synthesis, word stress is assigned automatically. The loss of stress that should affect words in some cases is often applied correctly, even if this cannot be guaranteed.

Most Norwegian dialects have two so-called word tones (*tonelag*, *tonem*) associated with the stressed syllable, referred to in Norwegian as *tonelag 1/enkelt tonelag/tonem 1* and *tonelag 2/dobbelt tonelag/tonem 2*. We here refer to them as (*word*) *accents: accent 1* and *accent 2*, respectively.

The synthesis is based on a Standard East Norwegian pronunciation like that of a representative speaker from the Oslo region. In this variety of Norwegian it is important to assign to words of more than one syllable the correct word accent. The text-to-speech system does this automatically as much as possible, although mistakes may occur, especially in many cases where the same spelling is used for two different words with different word accents. For instance, what is written *hender* is pronounced differently (with accent 1) when it means *hands* (the noun) and (with accent 2) when it means *happens*. The word *bønder* (*farmers*) is pronounced with accent 1, but *bønner* (*beans* or *prayers*) is pronounced with accent 2. The sentence *Det hender at bønder har bønner i sine hender* has the pattern *x 2x x 1x x 2x x xx 1x* where *x* symbolises an unstressed syllable, *1* a syllable with accent 1, and *2* a syllable with accent 2. The synthesis treats words of one syllable as having accent 1.

In compound words the stress usually lies on the first part of the compound (on that syllable which has the stress when this part is used as a word in its own right), but the accent often changes; particularly frequently, but by no means always, if the first part is a word of one syllable, the compound will have accent 2. In compound words, there is often a “reminder” of the original word stress in the last part; this is a kind of “lighter” stress, referred to as *secondary stress* or *secondary accent*. The synthesis often, but not always, correctly represents this secondary stress.

When writing phonetic text it is important to indicate the correct stress, especially the main stress (the word stress). It is important to realise that the word accents are an integrated aspect of the word stress. In other words, in order to indicate word stress one has to decide whether the word should be pronounced with accent 1 or accent 2. Put differently, indicating which word accent a syllable has implies that that syllable is stressed.

In the phonetic alphabet used in the text-to-speech system the word accents (and, implicitly, the word stress) are represented by a number placed immediately after the stressed vowel (no space character allowed in between), as follows (a space character must separate each symbol representing the speech sounds; an explanation of the symbols for consonants and vowels follows shortly):

Accent type	Accent symbol	Word	Transcription	Comment
accent 1	4	bønder	b 24 n @	
accent 1	4	finner	f i4 n @ r	verb
accent 2	3	bønner	b 23 n @	
accent 2	3	finner	f i3 n @ r	noun
secondary stress	1	bondegårder	b u4 n @ g O:1 r @ r	compound; secondary stress on the first syllable of the second part; word stress with accent 1 on the first syllable of the word
secondary stress	1	vannkanne	v A3 n k_h A1 n @	compound; secondary stress on the first syllable of the second part; word stress with accent 2 on the first syllable of the word

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.

7.2. Symbols for the Norwegian consonants

Table 7.1. Symbols for the Norwegian consonants

Symbol	Word	Phonetic text	Comment	RULSYS
b	be	b e:4		B
d	bade	b A:3 d @		D
rd	verdi	v { rd i:4	retroflex d	2D
f	far	f A:4 r		F
g	gå	g O:4		G
h	ha	h A:4		H
j	ja	j A:4		J
k	skal	s k A4 L		K
k_h	kan	k_h A4 n	aspirated k	Kh
l	sile	s i:3 l @		L
L	holte	h O3 L t @	dark L	LI
rl	jarle	j A:3 rl @	retroflex l	2L
rL	fæl	f {:4 rL	retroflex flap, "thick l"	
m	mor	m u:4 r		M
n	nord	n u:4 r		N
rn	barn	b A:4 rn	retroflex n	2N

Symbol	Word	Phonetic text	Comment	RULSYS
N	sang	s A4 N	not used syllable-initially	NG
p	spise	s p i:3 s @		P
p_h	pris	p_h r i:4 s	aspirated p	Ph
r	rar	r A:4 r		R
s	sa	s A:4		S
rs	vers	v {4 rs	retroflex s	2S
S	stasjon	s t A S u:4 n		SJ
t	stå	s t O:4		T
t_h	ta	t_h A:4	aspirated t	Th
rt	vert	v {4 rt	retroflex t	2T
rt_h	fortelle	f O rt_h e4 l @	aspirated retroflex t	Tr
C	kjole	C u:3 l @	not used syllable-finally	KJ
v	vi	v i:4		V
l=	handel	h A4 n d l=	syllabic l	
n=	femten	f e3 m t n=	syllabic n	
rn=	verten	v {4 rt rn=	syllabic rn	
aU	house	h aU s	used in e.g. English names	
x	Bach	b A x	German	X
w	Windows	w i4 n d O w s	English	W
tS	church	tS 2:4 tS	English	ch
dZ	John	dZ O4 n	English	
T	thing	T i4 N	English	th
D	this	D i4 s	English	dh
z	zoo	z u:4	English	Z
Z	Jean	Z A~4	French	

7.2.1. Comments on the phonetic symbols for consonants

7.2.1.1. 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 *bordell* and *modell*, *fort* and *fort*. In phonetic text these sounds are written as the combination of the *r* and the following dental consonant, with no space in between them. Note that not all dialects of Norwegian have retroflexes.

7.2.1.2. Aspiration

The voiceless stops *p*, *t*, *k* are typically aspirated in certain positions of the word. That is, they are followed by a “puff of breath” before the vowel. This happens when the stop sound is followed by a stressed vowel and in the beginning of words, including words that are part of compounds.

7.3. Vowels

Table 7.2. Symbols for the Norwegian vowels

Symbol	Word	Phonetic text	Comment	RULSYS
A	matt	m A4 t		A
A:	mat	m A:4 t		A:
e	vett	v e4 t		E
e:	vet	v e:4 t		E:
@	mate	m A:3 t @	schwa vowel allophone	E0
i	fille	f i3 l @		I
i:	file	f i:3 l @		I:
u	bonde	b u3 n @		O
u:	bone	b u:3 n @		O:
}	lutt	l }4 t		U
}:}	lut	l }:4 t		U:}
y	lynne	l y3 n @		Y
y:	lyne	l y:3 n @		Y:}
O	vått	v O4 t		Å
O:	våt	v O:4 t		Å:}
{	herr	h {4 r		Æ
{:	her	h {:4 r		Æ:}
2	møtt	m 24 t		Ø
2:	møt	m 2:4 t		Ø:}
Ai	hai	h Ai4	diphthong	AI
{}	haug	h {}4	diphthong	AU
{i	lei	l {i4	diphthong	EI
2y	køye	k_h 2y3 @	diphthong	ØY
A~	blanc	b l A~4	French, nasal A	
E~	vin	v E~4	French, nasal E	
O~	bon	b O~4	French, nasal O	
a:	Saar	z a:4 r	German A	

7.3.1. Comments to the phonetic symbols for vowels

7.3.1.1. Vowel length

Long vowels are marked with colon ':'.}

7.4. Glottal stops

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

7.5. Pause

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

Chapter 8. Abbreviations

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

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
ev.	eventuelt
evt.	eventuelt
resp.	respektive
tel.	telefon
tlf.	telefon
A.S.	Aksjeselskap (case sensitive)
AS	Aksjeselskap (case sensitive)
bl	blant
bl. a.	blant annet/andre (requires full stops)
bl.a.	blant annet/andre (requires full stops)
ca.	cirka
dvs.	det vil si
etc.	etcetera
jr.	junior
kl.	klokken
osv.	og så videre
m.m.	med mer
m.fl.	med flere
d.y.	den yngre
d.e.	den eldre

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.

Abbreviation	Reading
20 ml	tjue milliliter
25 cl	tjuefem centiliter
30 dl	tretti deciliter
40 mm	førti millimeter (<i>mm</i> is normally read as <i>med mer</i>)
50 cm	femti centimeter
60 dm	seksti decimeter
70 kg	sytti kilo
tel 32 12 34 56	telefon trettito (pause) tolv (pause) trettifire (pause) femtiseks
mob 245 35 043	mobil tohundreogførtifem (pause) trettifem (pause) null førtitre

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 *punktum*, hyphens '-' as *bindestreck*, underscores '_' as *understrekning*, slashes '/' as *skråstrek*.
- *no*, *uk*, *us* and all the other abbreviations for countries are spelled out letter by letter.
- The @ is read *krøllalfa*.
- 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

support@acapela-group.com

Reading

v v v punktum acapela bindestreck punktum com

support krøllalfa acapela bindestreck punktum com