

RealSpeak Solo (*) Escape Sequences

Many escape sequences listed below function as delimiters. When inserted in the middle of an input string, they delimit pieces of text to be separately processed and pronounced by the system. A consequence of inserting a delimiting escape sequence within an input string will be a short audible pause as well as an end-of-sentence intonation at the place of the escape sequence. To avoid this, it is recommended to insert such sequences only at natural boundaries, e.g. between sentences rather than within sentences.

The following escape sequences, however, can be inserted in the middle of an input string. They do not function as delimiters and therefore do not break the input into separate units. They are:

- <ESC>V (volume),
- <ESC>R (rate),
- <ESC>P (inserting pauses),
- <ESC>" (assigning sentence accent),
- <ESC>\pause=x\,
- <ESC>C (forcing a continuation),
- <ESC>/+ (entering L&H+ phonetic input),
- <ESC>@x (indicating part of speech),
- <ESC>\mrk=x\ (setting bookmarks),
- <ESC>\tn=x\.

* * *

<ESC>\vol=x\ : Selecting the volume

The <ESC>\vol=x\ setting determines the volume of the speech output.

Range (x) : 0..100 (0: silence - 100: maximum volume)

Default: 90

* * *

<ESC>\rate=x\ : Selecting the speech rate

The <ESC>\rate=x\ escape sequence determines how fast a text is spoken.

Range (x) : 1..100 (1: slow rate - 100: fastest rate)

Default: 50

* * *

<ESC>Mx : Selecting the read mode

This escape sequence determines the way in which the system will split the input text into message units. Each message unit will then be separately processed and pronounced by the Text-To-Speech system.

Four possibilities are offered:

Range (x): 0..3

0 The system will read character by character

1 The system will read the text word by word.

2 The system will read the text sentence by sentence. In order to divide a text into sentences, the system applies heuristics based upon the use of punctuation marks and capitalization.

3 The system will read line by line

Default: 2

NOTE

1. A distinction should be made between the input text provided to the Text-To-Speech system, and the actual message that is processed and pronounced by the Text-To-Speech system. Depending on the selected read mode, the input text is broken up into message units that are separately pronounced by the Text-To-Speech system.

* * *

<ESC>Wx : Setting a wait period between all succeeding message units

This control sequence generates additional pauses in between different Text-To-Speech message units.

Range (x) : 0..9

0 no additional pause

1 200 milliseconds

9 1800 milliseconds

Default: 2

NOTE

Depending on the read mode, message units can be:

1. Characters, words, sentences or terminator-delimited text blocks.
2. For double-byte languages, a message unit may be defined differently.

* * *

<ESC>@x : Indicating part-of-speech

This control sequence allows the user to specify the grammatical class of a word, in order to influence the pronunciation of that word. The control sequence is used in case of homographic heterophones, i.e. identical word forms that can be pronounced in different ways.

Range (x) :

N Noun; no additional pause

J Adjective

A Adverb

V Infinitive

R Past participle

v Verb, other forms

NOTE

1. For a considerable number of languages this control sequence is not supported or (its implementation) may deviate from this standard description.
2. The Text-To-Speech system will always give homographs their more frequent pronunciation. The control sequence <ESC>@x makes it possible to specify the part-of-speech to obtain the alternative pronunciation. You can also use the control sequence <ESC>@@, which allows you to get the alternative pronunciation of a homograph *without* specifying the part-of-speech. This is especially useful for words that have different pronunciations based on meaning rather than on part-of-speech.
3. Unlike all other control sequences, the part-of-speech control sequence is **case-sensitive**: the values N, J, A, V or R must be written in capitals, v (for infinitive) must be written in small letters.

* * *

<ESC>Px : Inserting pauses

This control will disappear in the following release. Please use

<ESC>|pause=x\ instead.

<ESC>|pause=x\ : Inserting long pauses

This control sequence inserts a pause *within* a message unit, as opposed to <ESC>Wx, which influences the pauses *in between* message units.

Range (x): 1..10,000,000

* * *

<ESC>" : Assigning sentence accent

Sentence accents are used to mark important words within a sentence.

In the RealSpeak system, they are assigned automatically.

The user can manipulate the way in which sentence accents are assigned, by using the control sequence <ESC>". The sequence <ESC>" is placed in front of the words requiring accentuation. However, manually inserting sentence accents may have no effect.

This depends on whether or not the system itself has already automatically assigned a sentence accent.

NOTE

For a number of languages this control sequence or its implementation may deviate from this standard description or it can be used but has no effect.

Refer to the section Quick Reference of the Control Sequences in

Chapter II: <Language> Text-To-Speech System of the User's Guide for <Language>.

* * *

<ESC>C : Forcing a continuation

This control sequence can be useful in sentence read mode (<ESC>M2) to prevent the system from inappropriately detecting the end of a sentence.

NOTE

For some languages this control sequence or its implementation may deviate from this standard description.

Refer to the section "Quick Reference of the Control Sequences" in **Chapter II: <Language> Text-To-Speech System of the User's Guide for <Language>.**

* * *

<ESC>E : Forcing an end-of-message

This control sequence can be seen as the counterpart of <ESC>C, discussed above. It forces an end-of-message in all read modes. It can be useful to customize the process of splitting a text into smaller units.

NOTE

For some languages this control sequence or its implementation may deviate from this standard description.

Please refer to the section **Quick Reference of the Control Sequences** in **Chapter II: <Language> Text-To-Speech System of the User's Guide for <Language>.**

* * *

<ESC>/+ : Entering L&H+ phonetic input

This control sequence allows the user to input text in phonetic transcription, using the L&H+ phonetic alphabet. This can be useful to ensure a correct pronunciation when words are mispronounced by the system, e.g. in case of foreign names. More information on this feature is given below in the section 'Using phonetic input'.

* * *

<ESC>%x : Setting the type of document

The control sequence <ESC>%x allows the user to switch from one preprocessing mode to another before pronouncing a text or generating PCM data. The two preprocessing modes that are currently available are standard text and e-mail. By default, the preprocessing mode is set to standard text.

Range (x) Text For standard text

E Mail For e-mail messages

Default Text

NOTE

For some languages this control sequence can be used but it has no effect.
Please refer to the section **Quick Reference of the Control Sequences** in **Chapter II: <Language> Text-To-Speech System** of the **User's Guide for <Language>**.

* * *

<ESC>\voice=x\ Selecting the voice

Where x = voice name

This control sequence allows the user to select one of the voices available for the specified language. This control sequence is not used to switch to the voice of different language.

NOTE

1. If the requested voice is not available, the currently loaded voice will not be changed.
2. Refer to the section **Quick Reference of the Control Sequences** in **Chapter II: <Language> Text-To-Speech System** of the **User's Guide for <Language>** for the available voices for a particular language and a particular technology.

* * *

<ESC>\Mrk=x\ : Inserting a bookmark

The control sequence <ESC>\Mrk=x\ allows the user to insert a bookmark in the text. The Text-To-Speech system will keep track of the marked position throughout the conversion of text into speech and will be able to locate the corresponding position in the speech signal. This can be useful in order to synchronize the input text with visual animation or to emphasize the part of the input text that is currently being spoken. The Text-To-Speech system supports a wide range of bookmark identity numbers.

Range (x): 0..4294967295

* * *

<ESC>\tn=x

The control sequence <ESC>\tn guides the TTS system on how to pronounce particular text normalization (tn) items that are ambiguous otherwise.

The following tn values are available:

<ESC>\tn=address

(This replaces the previous <ESC>\pron=address\)

This control sequence triggers the address-specific pronunciation mode. It ensures a pronunciation that is specific for addresses. By using <ESC>\tn=address\, the user can switch between standard reading mode and address mode.

When the address mode is on, the settings remain valid until explicitly reset by the reset sequence <ESC>F (reset default settings) or by setting <ESC>\tn=normal\.

<ESC>\tn=normal

This control sequence resets the default settings of the system. When the tn=normal mode is active, pronunciation is normal and not specific for address or spelling.

<ESC>\tn=spell

(This replaces the previous <ESC>\spell=on\)

This control sequence activates the spell mode and makes the TTS system start spelling out the input texts. This setting remains valid until explicitly reset by the reset sequence <ESC>F (reset default settings) or by setting <ESC>\tn=normal\.

* * *

<ESC>\prompt=x:y\ : Insert active prompt with identity 'y' from the domain 'x'

This control sequence inserts specified prompt "y" from the domain "x".

* * *

<ESC>\domain=x\ : Enable domain for automatic insertion

This control sequence enables specific domain to insert all prompts based on orthographical matching.

* * *

<ESC>F : Resetting default values

This control sequence resets the following control sequences to their default values:

<ESC>\vol=x\ volume 90

<ESC>\rate=x\ rate 50

<ESC>Mx read mode 2

<ESC>Wx wait period 2

<ESC>\tn=x\ normal

NOTE

As not all these control sequences are supported for some languages, some of them cannot be reset for those languages.

(*) RealSpeak Solo is a registered trade mark of Nuance Communications, Inc