Telecom RealSpeak / Host Software Development Kit

User's Guide for Brazilian Portuguese
V4.0

EVALUATION AGREEMENT

READ THE FOLLOWING CAREFULLY BEFORE USING THE SOFTWARE. PRIOR TO RECEIVING THE SOFTWARE, YOU HAVE SIGNED EITHER A LICENSE AGREEMENT OR A NON-DISCLOSURE AGREEMENT WITH SCANSOFT, CONCERNING THE REALSPEAK SOLO DEVELOPMENT SYSTEM OR SOME OF ITS COMPONENTS. THEREFORE, REFERENCE IS MADE TO THESE AND AS SUCH THE CONTRACTUAL TERMS AND CONDITIONS SHALL APPLY TO THE SOFTWARE.
IF YOU HAVE NOT SIGNED SUCH AN AGREEMENT AND IF YOU USE THE SOFTWARE, SCANSOFT WILL ASSUME THAT YOU AGREED TO BE BOUND BY THE EVALUATION AGREEMENT SPECIFIED HEREUNDER. IF YOU DO NOT ACCEPT THE TERMS OF THIS EVALUATION AGREEMENT, YOU MUST RETURN THE PACKAGE UNUSED TO SCANSOFT WITHIN SEVEN (7) DAYS AFTER RECEIPT.

1. Grant of Rights

In consideration of a possible commercial relationship, ScanSoft hereby grants to you, the LICENSEE, who accepts, a non-exclusive right to internally evaluate and test the software program ("the Software").

2. Ownership of Software

ScanSoft retains title, interests and ownership of the Software recorded on the original disk(s) and all subsequent copies of the Software and Documentation, regardless of the form or media in or on which the original and other copies may exist. ScanSoft reserves all rights not expressly granted to LICENSEE.

3. Copy Restrictions

This Software and the accompanying documentation are copyrighted. Unauthorized copying of the Software, including Software that has been merged or included with other software, or of the documentation is expressly forbidden. LICENSEE may be held legally responsible for any intellectual property infringement that is caused or encouraged by his failure to abide by the terms of this agreement. LICENSEE is allowed to make two (2) copies of the Software solely for backup purposes, provided that the copyright notice is included on the backup copy.

4. Use Restrictions

LICENSEE agrees not to use the Software for any other purpose than internally evaluating the Software. LICENSEE may physically transfer the Software from one computer to another, provided that the Software is used on only one computer at a time. LICENSEE may not modify, adapt, translate, reverse engineer, decompile, disassemble or create derivative works based on the Software.LICENSEE may not modify, adapt, translate or create derivative works based on the documentation provided by ScanSoft. The Software may not be transferred to anyone without the prior written consent of ScanSoft. In no event may LICENSEE transfer, assign, lease, sell or otherwise dispose of the Software and Documentation on a temporary or permanent basis except as expressly provided herein.

5. Warranty

THE SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ScanSoft shall have no liability to LICENSEE or any third party for any claim, loss or damage of any kind, including but not limited to lost profits, punitive, incidental, consequential or special damages, arising out of or in connection with the use or performance of the Software and accompanying documentation.

6. Termination

This agreement is effective until terminated. ScanSoft reserves the right to terminate this agreement automatically if any provision of this agreement is violated. LICENSEE may terminate this agreement by returning the Software and the accompanying documentation to ScanSoft, along with a written warranty stating that all copies have been returned.

Copyright

Telecom RealSpeak / Host
Software Development Kit
User's Guide for Brazilian Portuguese

Trademarks

Document History

Date	Type of change	SoftwareVersion
November 2004	Update for version 4.0	V4.0

Table of Contents

Telecom RealSpeak / Host SDK

Chapter 1

Brazilian Portuguese Text-To-Speech System

User's Guide for Brazilian Portuguese
V4.0

Chapter 1

Brazilian Portuguese Text-To-Speech System

Introduction

This section provides operational instructions for the Telecom RealSpeak /Host Text-To-Speech system for Brazilian Portuguese. It reviews the functionality of the system, and describes how the user can customize the pronunciation of input texts. This part also describes issues that are particular to the Brazilian Portuguese Text-To-Speech system. It introduces the Brazilian Portuguese phonetic alphabet and discusses some language-specific features of the Brazilian Portuguese Text-To-Speech system.

Preparing a text for Text-To-Speech

In general, there are four ways to intervene in the pronunciation of text:

- By using control sequences
- By entering phonetic input
- By using a user dictionary
- By using Microsoft SAPI5

Using Control Sequences

The following tables provide a reference for all supported control sequences in Brazilian Portuguese.

Remark: <ESC> represents the escape character "\x1B" (decimal 27) that generates the ASCII character 27 (Hex 1B).

Chapter 1

Quick Reference of the Control Sequences

Sequence	Description	Range	Default	Delimiter
<ESC> \vol=x	Volume (x : 0 .. 100)	0 = silence 10 = low 100 = high	80	No
	For example: <ESC>\vol=10\ Eu posso falar baixinho, <ESC>\vol=90\ mas também bem alto.			
<ESC> \rate=x	Speech Rate (x : 1 .. 100)	10 = slow 100 = fast	50	No
	For example: Eu posso <ESC>\rate=70\ acelerar minha fala <ESC>\rate=20\ ou então ralentá-la.			
<ESC> \rate_wpm=xxx	Words per minute (xxx: 1..1000)	Voice-specific (see subsequent table)	Voice-specific	No
	For example: Eu posso <ESC>\rate_wpm=350\ acelerar minha fala, <ESC>\rate_wpm=110\ ou então ralentá-la.			
<ESC>M x	Read Mode	0 = letter-by-letter 1 = word-by-word 2 = sentence-by-sentence 3 = line-by-line	2	yes
	For example: <ESC>M0 Demonstrador (The word " Demonstrador ".) <ESC>M1 Esse é o Demonstrador do Português Brasileiro (This sentence will be read word by word.) <ESC>M2 Esse é o Demonstrador do Português Brasileiro. (This sentence will be read as one sentence.)			
<ESC>W x	Wait Period	0 = no wait period 1 = 200 milliseconds wait period 9 = 1800 milliseconds	2	Yes
	For example: <ESC>M2 <ESC>W2 Heverá um curto período de silêncio após essa frase. <ESC>W9 Essa sentença será seguida por um longo período de espera. Você notou a diferença?			
<ESC> \Pause=xxx	Long Pause	1 .. 2^32-1 msec		no
	For example: Você pode inserir pausas entre as mensagens especificando <ESC>\pause=1500\ a sua duração.			
<ESC>"	Sentence Accent			no

Chapter 1

	For example: <ESC>"João vem amanhã. (i.e. not Pedro) João vem <ESC>"amanhã. (i.e. not today) Note: Manually inserted sentence accents may have no effect in RealSpeak. The RealSpeak synthesis module may indeed have reasons to override the requested sentence accent, and thus not realize it.			
<ESC>C	Continuation			no
	For example: Cidade das Flores. São Conrado Cidade das Flores.<ESC>C São Conrado In the first of the above examples, the text-to-speech system will detect an end-of-sentence afterFlores and hence pause before pronouncing the second part of the input. In order to make the system pronounce the entire input as one sentence, a continuation sequence should be inserted.			
<ESC>E	End-of-Message			yes
	For example: A Sra. Dra. Adelaide entrou na sala. Seja bemvinda, Sra. Dra<ESC>E Teve uma boa viagem? In the first of the above examples, the TTS system will appropriately read the input as one sentence. In the second example, the system would likewise read the two sentences of the input as one sentence. The sequence <ESC>E forces the system to read the entire input as two separate sentences.			
<ESC>/+	Phonetic Input (L&H+ phonetic alphabet)			no
	For Example: <ESC>/+pe.'Zo<ESC>/+			
<ESC>%x	Preprocessing Mode	s = string text = standard text mode email = e-mail mode		yes
	For example: <ESC>%text Author: Pedro Santos. <ESC>%email Author: Pedro Santos			
<ESC>\tn=s	Guide text normalization	s = string: address, spell, normal		
	for example: <ESC>\tn=address\ rod. Bandeirantes, S. André, cep 09089-678 <ESC>\tn=normal\ <ESC>\tn=spell\ soletrar.<ESC>\tn=normal\ soletrar.			
<ESC>F	Reset to Default			yes

Chapter 1

	For example: <ESC>\vol=100\ O volume mais alto. <ESC>F O volume padrão. <ESC>\rate=100\ A velocidade mais alta. <ESC>F A velocidade padrão.

Sequence	Description	Range	Default	Delimiter
<ESC>@c	Declare the part-of-speech	c = character		No
<ESC>\domain=s\	Enable the extension	s = string: the name of the extension		Yes
<ESC>\domain\	Disable the last extension			Yes
<ESC>\voice=s\	Set the voice	s = string: the name of the voice		Yes
<ESC>\mrk=n\	Insert a bookmark	n = 0..4294967295		No
<ESC>\p\	Insert a paragraph boundary			Yes
<ESC>\audio=”s“\	Insert an audio file	s = string: the filename URI		No

Chapter 1

Speech Rates in Words per Minute for Brazilian Portuguese Voices

Words per minute		
Voice	**Range**	**Default**
Raquel	Min = 75 Max = 377	151

Entering phonetic input

How to proceed

To switch from orthographic to phonetic mode, insert <ESC>/+ to use the L&H+ phonetic alphabet. The phonetic input mode remains active until the command is explicitly reset by entering <ESC>/+ again.

The phonetic input string is composed of symbols of the L&H+ phonetic alphabet (see phonetic table below). Examples are given in the phonetic table below.

Chapter 1

In addition to the phonetic symbols, it is advised to use the following characters in the phonetic input string:

Special characters		
L&H + Symbol	**Meaning**	**As in:**
' (ASCII 39, Hex 27)	primary word stress	<ESC>/+ fa.'la.r6a0<ESC>/+ (falara) vs. <ESC>/+ fa.la.'r6a<ESC>/+ (falará)
'2	secondary word stress	<ESC>/+ ka.'2fE.'zi%~.j%~u<ESC>/+ (cafezinho)
" (ASCII 34, Hex 22)	sentence accent	<ESC>/+"nE.sta0_'fr6a.zi_a_u% ~_a.'se%~.tu <ESC>/+ (nessa frase há um acento)
.	syllable boundary	<ESC>/+ 'si.la.ba0 <ESC>/+ (sílaba)
#	silence (pause)	<ESC>/+'na%~w%~_"sej#'a.Su _ki_"si%~<ESC>/+ (Não sei...Acho que sim.)

Note that the use of punctuation marks remains useful within phonetic input to assure a correct intonation. Each punctuation mark needs to be preceded by an asterisk.

For example:

```
<ESC>/+
va.'lew_a_"pe%~.na0*?#'tu.du_'va.li_a_"pe%~.n
a0*,_si_a_"aw.ma0_'na%~w%~_'E_pe."ke%~.na0*!<
ESC>/+
```

(Valeu a pena? Tudo vale a pena se a alma não é pequena.)

Chapter 1

Punctuation Marks	
L&H+ Symbol	**Meaning**
_	Word delimiter
*.	End of declarative
*,	Comma
*!	End of exclamation
*?	End of question
*;	Semicolon
*:	Colon

Lexical stress and sentence accents in phonetic input

In phonetic input strings, `lexical stress` and `sentence accents` can be manually indicated by the user, by using a single quote (') or double quote (") respectively.

Note that manually inserted lexical stress or sentence accents may have no effect in RealSpeak. The RealSpeak synthesis module may indeed have reasons to override the requested sentence accent, and thus not realize it.

- The Text-To-Speech system will automatically convert all lexical stress marks into sentence accents in case no manually added sentence accents are found in the phonetic input string.

  For example:

  ```
  <ESC>/+u_'e.zi.tu_du_pr6o.'ZE.tu_ta%~.'be%~
  j%~_de.'pe%~.d&Zi_d&Zi_'nOs*.<ESC>/+
  ```
  (O êxito do projeto também depende de nós.)

  is the same as:

  ```
  <ESC>/+u_"e.zi.
  tu_du_pr6o."ZE.tu_ta%~."be%~j%~_de."pe%~.d&
  Zi_d&Zi_"nOs*.<ESC>/+
  ```

Chapter 1

- If phonetic input contains at least one manually added `sentence accent`, no additional sentence accents are assigned by the text-to-speech system. Therefore, only those words marked with " will get a sentence accent. As a consequence, a message containing only one manual sentence accent will have an almost flat intonation on all the other words.

  For example:

  ```
  <ESC>/+u_'e.zi.tu_du_pr6o.'ZE.tu_ta%~."
  be%~j%~_de.'pe%~.d&Zi_d&Zi_'nOs*.<ESC>/
  +
  ```

  (Only one sentence accent will be realized.)

- Phonetic input can also be combined with orthographic input.
  If no sentence accents are found in the input text (indicated by <ESC>" in orthographic input, or by " in phonetic input), the Text-To-Speech system will automatically assign sentence accents. In the orthographic part of the input, the Text-To-Speech system will realize these sentence accents on the basis of part-of-speech and syntactic information. In the phonetic part of the input, all lexical stress marks (if any) will be converted into sentence accents. If there are no lexical stress marks, no sentence accent will be realized for the phonetic part of the input (see point 1 above).
  If the user has manually specified one or more sentence accents, no additional sentence accents will be realized (see point 2 above).

Chapter 1

For example:

```
O <ESC>/+ 'e.zi.tu <ESC>/+ do projeto
também depende de nós.
```

(No sentence accents are found; the Text-To-Speech system will automatically assign sentence accents.)

```
O <ESC>/+ "e.zi.tu <ESC>/+ do projeto
também depende de nós.
```

(A sentence accent is specified in the phonetic part of the input text. No additional sentence accents will be realized.)

```
O <ESC>/+ 'e.zi.tu <ESC>/+ do projeto
<ESC>"também depende de nós.
```

(A sentence accent is specified in the orthographic part of the input text. No additional sentence accents will be realized.)

```
O <ESC>/+ "e.zi.tu <ESC>/+ do projeto
<ESC>"também depende de nós.
```

(Two sentence accents were specified; no additional sentence accents will be realized.)

Chapter 1

The Brazilian Portuguese L&H+ and UNIPA Phonetic Alphabets

Vowels				
L&H+ Symbol	**L&H+ Transcription**	**UNIPA Symbol**	**UNIPA Transcription**	**As in:**
a	‘g**a**.tu	a	‘g**a**.tu	gato
a0	'pE.dr6**a0**	a0	'pE.dr6**a0**	pedra
a%~	‘k**a%~**. pu	a%~	‘k**a%~**. pu	campo
E	‘p**E**	E	‘p**E**	pé
e	‘s**e**.du	e	‘s**e**.du	cedo
e%~	‘t**e%~**.pu	e%~	‘t**e%~**.pu	tempo
i	't&Si.pu	i	't+Si.pu	tipo
i%~	‘l**i%~**. pu	i%~	‘l**i%~**. pu	limpo
i:~	p**i:**~.'new	i:~	p**i:**~.'new	pneu
O	‘p**O**	O	‘p**O**	pó
o	‘b**o**.lu	o	‘b**o**.lu	bolo
o%~	‘t**o%~**.tu	o%~	‘t**o%~**.tu	tonto
u	‘p**u**.la0	u	‘p**u**.la0	pula
u%~	‘t**u%~**.ba0	u%~	‘t**u%~**.ba0	tumba

Chapter 1

Consonants				
L&H+ Symbol	L&H+ Transcription	UNIPA Symbol	UNIPA Transcription	As in:
Z	‘**Z**e.lu	Z	‘**Z**e.lu	gelo
S	‘**S**a	S	‘**S**a	chá
h	poh.‘ke	h	poh.‘ke	porque
f	‘**f**a.ka0	f	‘**f**a.ka0	faca
v	‘va.li	v	‘va.li	vale
s	‘**s**e.du	s	‘**s**e.du	cedo
z	a.‘**z**ah	z	a.‘**z**ah	azar
p	‘**p**o.vu	p	‘**p**o.vu	povo
b	‘**b**O.la0	b	‘**b**O.la0	bola
t	‘**t**aw	t	‘**t**aw	tal
d	‘**d**a.du	d	‘**d**a.du	dado
k	‘**k**a.za0	k	‘**k**a.za0	casa
g	‘**g**a.tu	g	‘**g**a.tu	gato
m	‘**m**a.la0	m	‘**m**a.la0	mala
n	‘a%~.**n**u	n	‘a%~.**n**u	ano
n~	‘**n~**a%~.du	n~	‘**n~**a%~.du	nhando
l	‘**l**a.du	l	‘**l**a.du	lado
l~	‘fi.**l~**u	l~	‘fi.**l~**u	filho
r6	‘ka.**r6**u	r6	‘ka.**r6**u	caro
k&s	i.‘nO.**k&s**i:~	k+s	i.‘nO.**k+s**i:~	inox
d&Z	'**d&Z**i.a0	d+Z	'**d+Z**i.a0	dia
t&S	'**t&S**i.a0	t+S	'**t+S**i.a0	tia
j	‘pa**j**	j	‘pa**j**	pai
w	‘ka**w**.ma0	w	‘ka**w**.ma	calma
j%~	'te%~**j%~**	j%~	'te%~**j%~**	tem
w%~	'ma%~**w%~**	w%~	'ma%~**w%~**	mão

Chapter 1

NOTES

Note that the L&H+ alphabet is not SSML compliant. For SSML, use the UNIPA alphabet.

Using a User Dictionary

For information on how to create and use user dictionaries, please refer to Appendix C in the Telecom RealSpeak/Host Programmer's Guide.

Chapter 1

Using the Microsoft SAPI5 Lexicon

Microsoft SAPI5 provides lexicons so that users and applications can specify pronunciation and part-of-speech information for particular words. As such, all SAPI compliant Text-To-Speech engines should use these lexicons to guarantee uniformity of pronunciation and part of speech information.

There are two types of lexicons in SAPI: user lexicons and application lexicons.

User Lexicons

Each user who logs in to a computer will have a User Lexicon. Initially, this lexicon is empty; words can be added either programmatically, or by using an engine's add/remove words UI component (for example, the sample application Dictation Pad provides an Add/Remove Words dialog).

Application Lexicons

Applications can create and ship their own lexicons of specialized words. These lexicons are fixed and cannot be edited.

Detailed information on how to use the MS SAPI5 lexicons can be found in the manual "Microsoft Speech SDK V5.1", chapter "ISpLexicon Interface".

The Brazilian Portuguese SAPI5 Phoneme List

To add entries to the lexicon, the user should use a set of language specific phonemes. The language specific phoneme list for Brazilian Portuguese is given below.

Chapter 1

SAPI5 Symbols			
SAPI5 Symbol	**As in:**	**SAPI5 Transcription**	**SAPI Phone ID**
A	gato	S1 G A . T U	0061
A nas	campo	S1 K A nas . P U	0061 0303
AEX	pedra	S1 P EH . D DX AEX	0250
EH	pé	S1 P EH	025B
E	cedo	S1 S E . D U	0065
E nas	tempo	S1 T E nas . P U	0065 0303
I	tipo	S1 T + SH I . P U	0069
I nas	limpo	S1 L I nas . P U	0069 0303
I xsh	rock	S1 H AO . K I xsh	0069 02D8
AO	pó	S1 P AO	0254
O	bolo	S1 B O . L U	006F
O nas	tonto	S1 T O nas . T U	006F 0303
U	pula	S1 P U . L AEX	0075
U nas	tumba	S1 T U nas . B AEX	0075 0303
ZH	gelo	S1 ZH E . L U	0292
SH	chá	S1 SH A	0283
H	porque	P O H . S1 K E	0068
F	faca	S1 F A . K AEX	0066
V	vale	S1 V A . L I	0076
S	cedo	S1 S E . D U	0073
Z	azar	A . S1 Z A H	007A
P	povo	S1 P O . V U	0070
B	bola	S1 B AO . L AEX	0062
T	tal	S1 T A W	0074
D	dado	S1 D A . D U	0064
K	casa	S1 K A . Z AEX	006B

Chapter 1

SAPI5 Symbols			
SAPI5 Symbol	**As in:**	**SAPI5 Transcription**	**SAPI Phone ID**
G	gato	S1 G A . T U	0067
M	mala	S1 M A . L AEX	006D
N	ano	S1 A nas . N U	006E
NJ	nhando	S1 NJ A nas . D U	0272
L	lado	S1 L A . D U	006C
LJ	filho	S1 F I . LJ U	028E
DX	caro	S1 K A . DX U	027E
K + S	sexo	S1 S EH . K + S U	0068 0361 0073
J	pai	S1 P A J	006A
W	calma	S1 K A W . M AEX	0077
D + ZH	dia	S1 D + ZH I . AEX	0064 0361 0292
T + SH	tia	S1 T + SH I . AEX	0074 0361 0283
J nas	tem	S1 T E nas J nas	006A 0303
W nas	m□	S1 M A nas W nas	0077 0303

Chapter 1

Notes on the Brazilian Portuguese Text-To-Speech System

The Brazilian Portuguese Text-To-Speech system has been designed to allow a correct pronunciation of any input written according to the rules of Brazilian Portuguese orthography. The following cases, however, require special attention.

Cardinal Numbers

Cardinal numbers up to 15 digits are pronounced as full numbers. Periods may be used to separate groups of digits. Digit strings consisting of more than 15 digits are pronounced digit by digit.

For example:

```
1234567
1.666
007897989465736223
```

Decimal Numbers

Decimal numbers may consist of up to 15 digits before and after the comma. If one or two digits occur after the comma, they will be read as full numbers; if more than 2 digits appear after the comma, they will be spelled.

For example:

```
22,30
0,25800
```

Chapter 1

Ordinal Numbers

The following numbers are pronounced as ordinal numbers:

- Cardinal numbers from 1 to 999 followed by ª or °
- Cardinal numbers starting with a digit other than zero, followed by 1 to 14 zeros and followed by ª or °. Periods can divide the digit string into groups of 3 digits.

For example:

```
2°
134°
```

Roman Numbers

Roman numbers are pronounced as ordinal or cardinal numbers, depending on the context. E.g. if a name precedes a roman number, the roman number will be read as an ordinal, preceded by "the". If the roman number is preceded by an abbreviation like "p.", it is pronounced as a cardinal. Roman numbers consisting of only one character, without a valid context, are not considered as roman numbers.

For example:

```
Rainha Catarina II      Rainha Catarina
                        Segunda
capítulo XX             capítulo vigésimo
século XXI              século vinte e um
```

Telephone Numbers

The Brazilian Portuguese system supports telephone numbers in the following formats:

- with operator code (optional) and country code
- with operator code (optional) and area code
- without area code preceded by one of the following abbreviations: t, tel, tele, telef, fone, fax.

The operator code, the country code and the area code may be surrounded by parentheses or may be separated by means of a hyphen or a space. You can also use a hyphen or a space to separate the last four digits of the telephone number. You cannot change the type of separator in one and the same telephone number.

Chapter 1

For example:

```
tel. 6941 2786
fone 550-2812
011 6942 5781
(015) (32) 950-2742
0015-1-212-555-7781
```

Dates

The Brazilian Portuguese Text-To-Speech system reads dates in the following formats:

Day (1 or 2 digits)/Month (1 or 2 digits)/Year (2 to 4 digits)
Day (1 or 2 digits)-Month (1 or 2 digits)-Year (2 to 4 digits)

If the year is indicated by 2 digits, it may be preceded by a quote.

For example:

```
08-02-'95
16/05/1991
16.05.00
```

Dates can also be specified in written format:

For example:

```
8 jul. 68
30 ABRIL 1999
```

Chapter 1

Time Indications

Time indications will be correctly pronounced when written in one of the following formats:

hour (1 or 2 digits) followed by "h" or "hs"

```
06h    seis horas
12hs   doze horas
```

hour (1 or 2 digits) followed by "h" and minutes (2 digits) optionally followed by "min"

```
6h32min            seis horas e trinta e dois
minutos
23h10min     vinte e três horas e dez minutos
```

hour (1 or 2 digits) and minutes (2 digits) separated by colon or period, followed by "h", "hs", "am" or "pm"

```
2:30 min           duas horas e trinta minutos
10.30pm            dez horas e trinta minutos
pm
12:00hs            doze horas e zero minutos
```

hour (1 or 2 digits) and minutes (2 digits) separated by a colon

```
23:55        vinte e três e cinquenta e cinco
8:30         oito e trinta
```

hour (1 or 2 digits), minutes (2 digits) and seconds (2 digits) separated by a colon

```
21:00:32           vinte e uma horas zero
minutos e trinta e dois segundos
9:23:43            nove horas vinte e três
minutos e quarenta e três
             segundos
```

Currencies

The Brazilian Portuguese Text-To-Speech system correctly handles the currency symbols R$, Cr$, NCr$, Cz$, US$, $ G/, Ur$, EURO, EUR, € and ¥. The currency sign should precede the numeral.

For example:

```
US$40              quarenta dólares
R$30,25            trinta reais e vinte e
                   cinco centavos
Cr$3.000,00        três mil cruzeiros
```

The Brazilian Portuguese TTS system also supports currencies in the following formats:

Chapter 1

- currency indicator + 1 to 6 digits (optionally followed by 1 or 2 digits after the comma) + number indicator: milhão, bilhão, trilhão, mi, bi, and tri.

  For example:

  ```
  R$23 milhões
  US$ 10,5 bi
  ```

Other currencies are written in full words and have to follow the numeral.

For example:

```
100 francos
20 liras
```

Abbreviations

The Brazilian Portuguese Text-To-Speech system contains a dictionary with the most common abbreviations, such as: `dr.`, `ex.`, etc.

Some abbreviations are ambiguous, however, and are pronounced depending on the context in which they appear. For example, the abbreviation km is pronounced "quilômetro" when preceded by the number 1, but "quilômetros" when preceded by a number other than 1. Similarly, cap. is pronounced "capítulo" when followed by a number, but "capitão" in all other contexts.

For example:

```
21 km           vinte e um quilômetros
1 km            um quilômetro.
cap. 4          capítulo quatro
cap. Souza      capitão  Souza
```

Acronyms

Words such as HIV, USP, PT, UNICEF, etc. are pronounced using a combination of rules and dictionaries.

```
HIV and PT are spelled, while USP and
UNICEF are pronounced as words.
```

Telecom RealSpeak / Host SDK

Chapter II

E-Mail Preprocessor

User's Guide for Brazilian Portuguese
V4.0

Chapter II

E-Mail Preprocessor

Introduction

The ScanSoft e-mail preprocessor (EMPP) has been developed to analyze a specific type of text: e-mail messages. E-mail messages differ from any average type of text in both structure and contents.

An e-mail message consists of two clearly distinguished parts: the header and the body. A substantial part of the header contains routing and administrative information, which is irrelevant to the user. Both the header and the body contain all kinds of e-mail specific text features, e.g. e-mail addresses, emoticons such as smileys, etc. Furthermore, informal writing is often combined with a lack of grammatical conventions. Spelling rules are frequently violated, punctuation is often omitted, etc.

Although the standard ScanSoft Text-To-Speech system can handle special text items (abbreviations, numbers, dates, etc.), it is not capable of correctly handling all e-mail specific text features. These text features are therefore dealt with by the e-mail preprocessor. The EMPP transforms e-mail specific information into a format that complies with the rules of the standard ScanSoft Text-To-Speech system. The EMPP is a plug-in preprocessing module of the ScanSoft Text-To-Speech system. It replaces the preprocessor of the standard Text-To-Speech system.

In the following sections you will find a description of the functioning of the ScanSoft e-mail preprocessor as well as an overview of its features.

The e-mail preprocessor has two main tasks: processing of the e-mail header and processing of the body of the e-mail message.

The input to the EMPP consists of one or more e-mail messages. In order to process the e-mail header, the EMPP extracts relevant header fields and then provides an intelligent header field reading.

Chapter II

During the processing of the e-mail body, the text is divided into smaller text units, called text-to-speech messages, which are synthesized by the Text-To-Speech system. Text normalization is applied to e-mail specific text features such as e-mail addresses, proper names, emoticons, URLs (Universal Resource Locators), etc. For the text normalization of an e-mail message, the ScanSoft EMPP applies linguistic rules and performs dictionary look-up, in order to yield an adequate phonetic transcription. The EMPP also supports the ScanSoft user dictionary mechanism, which allows the user to customize the output of the e-mail processing.

E-Mail Header Processing

Header Field Extraction

An e-mail message consists of two clearly distinguished parts: the header and the body. The EMPP detects the header and extracts the relevant header fields. Information that is of no interest to the user (such as routing information) is not retained.
The EMPP extracts the following header fields:

From Field	Contains the sender's name and/or address
Date Field	Contains the date and time of sending
Subject Field	Optionally contains the subject of the e-mail

The extraction of the header fields is based on the detection of specific keywords in the e-mail header. The supported keywords for the extraction of the header fields are listed below:

From Field	From: Author: Sender: De: Von:
Date Field	Date: Enviado: Gesendet:
Subject Field:	Subject: Subj: Asunto: Betreff:

Chapter II

The following is an example of header field extraction. The original header holds information that is irrelevant to the user. After extraction of date, sender and subject, the processed header merely mentions the Date field, the From field and the Subject field:

Original header:

Date: Tue, 7 Sep 1999 11:11:02 -0300 (EST)
Message-Id: <199909071411.LAA06415@programadoresde.com.br>
To: Carla.Strobli@scansoft.com
Subject: Bem-vindo a lista PER
Reply-To: majordomo@programa.com.br
SMTPOriginator: owner@programa.com.br
From: majordomo@programa.com.br
PostedDate: 09/07/99 04:11:02 PM
SendTo: CN=Carla
Strobil/OU=LIE/OU=BE/O=SCANSOFT@SCANSOFT
ReplyTo: majordomo@programa.com.br
$MessageStorage: 0
$UpdatedBy:
CN=LNMAILMTA01/OU=SERVER/OU=BE/O=SCANSOFT
RouteServers:
CN=LNMAILMTA01/OU=SERVER/OU=BE/O=SCANSOFT,C
N=LNMAILIEP01/OU=SERVER/OU=BE/O=SCANSOFT
RouteTimes: 09/07/99 04:29:50 PM-09/07/99 04:29:51
PM,09/07/99 04:11:11 PM-09/07/99 04:11:12 PM
$Orig:

DeliveredDate: 09/07/99 04:11:12 PM
Categories:
$Revisions:

Extracted header fields:

Date: Tue, 7 Sep 1999 11:11:02 -0300 (EST)
Subject: Bem-vindo a lista PER
From: majordomo@programa.com.br

Chapter II

Header Field Reading

After the header fields have been extracted, they are processed by the EMPP. The header field keywords (see above) are replaced by an introductory message. The remainder of the header fields is processed by the EMPP in order to allow the Text-To-Speech system to intelligently read the fields.

From Field

The **From** field keyword is replaced by the introductory message "`Mensagem de:`"

For example:

```
Author: Pedro Gabeira
```

is pronounced:

```
Mensagem de: Pedro Gabeira
```

The remainder of the **From** field is further processed by the EMPP. The EMPP supports **From** fields that either consist of

a) a proper name
b) a proper name and an address
c) an address

a) - b) In case the **From** field contains a proper name, this name and only this name is sent to the Text-To-Speech system. This means that if both a name and an address are found in the **From** field, the address will not be read by the Text-To-Speech system.

For example:

```
Author: Alexandre Silva
From: (João Souza) souza@scansoft.com
From: Anaidra@uunet.be (Adriana Antunes)
```

are pronounced:

```
Mensagem de: Alexandre Silva
Mensagem de: João Souza
Mensagem de: Adriana Antunes
```

Chapter II

c) In case the **From** field contains only an address, the EMPP extracts the name out of the address and reads the domain name literally.

For example:

```
Author: meluna@ss.org.br
From: erica@scansoft.com
```

are pronounced:

```
Mensagem de: meluna arrôba s s ponto o r g
ponto b r
Mensagem de: erica arrôba scansoft ponto com
```

Date Field

The **Date** field keyword is replaced by the introductory message "`Date:`".

The **Date** field contains the date and time of sending. The EMPP supports multiple date and time formats, which are transformed into a uniform format that complies with the rules for date and time indications of the ScanSoft Text-To-Speech system. The EMPP only pronounces the date.

The EMPP supports dates in the following formats:

For example:

```
Date: 16 Oct 1996 22:37:02 +0100
Date: Segunda-feira, 30 de Agosto de 1999
17:26
Date: Sat, 21 Sep 2000 23:45:00 GMT
```

are pronounced:

```
Data: 16 do dez de 1996
Data: Segunda-feira, 30 de Agosto de 1999
Data: Sábado, 21 do nove de 2000.
```

Chapter II

Subject Field

The **Subject** field keyword is replaced by the introductory message "Assunto".

The **Subject** field can contain all kinds of data, but may also be empty. The EMPP searches for keywords that are typical for the subject field (e.g. RE, FYI, FW).

For example:

```
Subject: RE: sua chegada
Subject: FYI: ano letivo
Subject: FW: natal com neve
Subject: EN: pesquisas de opinião
```

are pronounced:

```
Assunto: Resposta: sua chegada
Assunto: Para sua informação: ano letivo
Assunto: Mensagem encaminhada: natal com neve
Assunto: Mensagem encaminhada: pesquisas de
opinião
```

Chapter II

E-Mail body processing

Message Extraction

The e-mail preprocessor splits the body of the e-mail message into text-to-speech messages. This is done on the basis of a number of criteria, such as punctuation, capitalization, layout, intelligent abbreviation handling, etc.

The following examples illustrate some criteria for splitting the e-mail text into text-to-speech messages:

- Using sentence final punctuation and capital letters

  ```
  Se você tiver alguma informação, por favor
  entre em contato. Escreva para o endereço
  ania@scansoft.com. Obrigada!
  ```

- Using layout

  ```
  Agenda da semana:
  1) Almoço com novos clientes
  2) Relatório de Progresso Mensal
  3) Cotações da semana
  ```

- Using intelligent abbreviation handling

  ```
  Para maiores informações, entrar em contato
  com M.S. Oliveira.
  ```

Chapter II

Text Normalization

An e-mail message typically contains e-mail specific text features, such as e-mail addresses, URLs, file names, emoticons, etc. The EMPP transforms these e-mail specific features into a format that complies with the rules of the standard text normalization of the ScanSoft Text-To-Speech system.

The following are examples of e-mail specific text normalization:

- Support for multiple e-mail address formats

  ```
  iane@penta.cesup.br
  Vaz@p4.f3.n2.z1.fidonet.org
  help@anon.penet.be
  ```

- Support for URLs (Universal Resource Locators)

  ```
  gopher://www.lycos.com
  http://www.serv.org/voluntarios
  telnet://www.empregos.com
  ```

- Support for file names

  ```
  ldb001.tse
  sysinfo.exe
  lipedu.xls
  ```

- Processing of emoticons

  ```
  :-*  is pronounced     "smack"
   ;-)              "ri ri ri"
   :-D              "ra ra ra"
  ```

- Processing of overuse of punctuation

  ```
  Cuidado!!!!!!!Vírus!!!!!!!!!
  Socorro, tem um bug #§#>>§< no meu arquivo!
  ```

  becomes:

  ```
  Cuidado! Vírus!
  Socorro, tem um bug no meu arquivo!
  ```

Chapter II

- Processing of Question/Answer (FAQ)

```
Q. I have an e-mail address change. How can
I ensure that I will continue to receive my
EcoLink e-mail newsletter?
A. Easy. Just send an e-mail to
EcoLink@peach.ease.msoft.com. Be sure to
include both your old address and your new
address.
```

becomes:

```
Question:
I have an e-mail address change. How can I
ensure that I will continue to receive my
EcoLink e-mail newsletter?
Answer:
Easy. Just send an e-mail to
EcoLink@peach.ease.msoft.com. Be sure to
include both your old address and your new
address.
```

- Processing of inserted mail

```
Paulo> Ana, ONDE ESTÁ VOCÊ? Por favor, não
Paulo> vá embora! Sei que você ainda
Paulo> está aí, responda por favor!
Ana> Tudo bem, aqui estou de volta.
```

becomes:

```
Paulo:
Ana, ONDE ESTÁ VOCÊ? Por favor, não
 vá embora! Sei que você ainda  está aí,
responda por favor!

Ana:
Tudo bem, aqui estou de volta.
```

Chapter II

Customizing the E-Mail Preprocessor

The e-mail preprocessor supports the standard ScanSoft Text-To-Speech SDK user dictionary mechanism, which allows the user to customize the output of the e-mail preprocessor. The user dictionary is consulted both during the header processing and the body processing.

For more information on how to build and use user dictionaries, see **Appendix C: User Dictionaries** of the `Programmer's Guide`.

Customizing the E-Mail Header

The user dictionary is consulted during the header processing while reading the **From** field and the **Subject** field.

From Field

The **From** field either consists of

a) a proper name
b) a proper name and an address
c) an address

a) In case the **From** field contains a proper name only, the name is passed on to the user dictionary. If the lookup is successful, the proper name is substituted by the replacement string. If not, the name is further processed by the header reading module.

For example:

If the user dictionary contains the following line:

```
Johan            /+'jo.ha%~
```

the following From field:

```
From: Johan Meyer
```

becomes:

```
Mensagem de: */+'jo.ha%~*/+ Meyer
```

Chapter II

b) In case the **From** field contains a proper name and an address, the EMPP first passes the address to the user dictionary. If the lookup is successful, both the proper name and the address are substituted by the replacement string. If not, the EMPP passes the proper name to the user dictionary. If this lookup is successful, the name and the address are substituted by the replacement string. If not, the name is further processed by the header reading module. The address will not be read by the Text-To-Speech system.

For example:

If the user dictionary contains the following lines:

```
silva@scansoft.com    Pedro, meu companheiro de tênis
Jonas                 meu melhor amigo
Gates                 /+'gej.t&Sis
```

the following **From** fields:

```
From: "Pedro Silva" silva@scansoft.com
Author:  jonas@elis.rug.ac.be (Jonas)
From: Gates@scansoft.com (J. Gates)
```

become:

```
Mensagem de: Pedro, meu companheiro de tênis
Mensagem de: meu melhor amigo
Mensagem de: J. */+'gej.t&Sis*/+
```

c) In case the From field contains only an address, the complete address is looked up in the user dictionary. If the lookup is successful, a proper name is added to the From field. If not, only the domain part is sent to the user dictionary. The EMPP first calls the dictionary for the complete domain part. If the lookup is successful, the complete domain part is substituted by the replacement string. Otherwise, the EMPP cuts off the leftmost sublevel domain and repeats the lookup and matching procedure for the remainder of the domain part. If the lookup is successful, the remainder of the domain part is substituted by the replacement string. This procedure is called repeatedly until the top level domain is encountered. If none of the lookups is successful, the address is further processed by the header reading module.

Chapter II

For example:

If the user dictionary contains the following lines:

```
jose.almeida@gsp.br            grande amigo
```

the following From field:

```
Sender: jose.almeida@gsp.br
```

becomes:

```
Mensagem de: grande amigo
```

NOTES

To allow a correct processing of the From field, the replacement string in the user dictionary should not contain an address or a domain.

Chapter II

Subject Field

Each word in the **Subject** field is sent to the user dictionary. If the lookup is successful, the replacement string is sent directly to the Text-To-Speech system. If not, the **Subject** field is further processed by the header reading module.

For example:

If the user dictionary contains the following lines:

```
ECAD      /+e.se.'a.de
IDT       I D T
```

the following Subject fields:

```
Subject: ECAD Relações Humanas
Subject: códigos para IDT
```

are pronounced:

```
Assunto: */+e.se.'a.de */+ Relações Humanas
Assunto: códigos para I D T
```

Chapter II

Customizing the E-Mail Body

When a user dictionary has been loaded, the EMPP will call the dictionary for every word of the e-mail body. If the word is found in the user dictionary, it is substituted by the replacement string. If not, the body is further processed by the e-mail body processing module.

For example:

If the user dictionary contains the following line:

```
lgtca   linguística
```

the word " `lgtca` " in the following sentence:

```
Eu não sou do departamento de lgtca mas
gostaria de participar da discussão.
```

is replaced by the corresponding string found in the e-mail user dictionary:

```
Eu não sou do departamento de lingüística mas
gostaria de participar da discussão.
```

Telecom RealSpeak / Host SDK

Chapter III

SSML Preprocessor

User's Guide for Brazilian Portuguese
V4.0

Chapter III

SSML Preprocessor

Introduction

SSML (Speech Synthesizer Markup Language) is part of a set of markup specifications for voice browsers. SSML was designed to provide a rich, XML-based markup language for assisting the generation of synthetic speech in web and other applications. The essential role of the markup language is to provide a standard way to control aspects of speech such as pronunciation, volume, pitch, rate.

The Telecom RealSpeak/Host SDK provides a preprocessor that supports a subset of the Speech Synthesizer Markup Language (SSML) specification. That subset is called "ScanSoft SSML" (4SML).

More information on SSML can be obtained from the "Speech Synthesis Markup Language specification – W3C Working Draft 5 April 2002" available at the W3C web site. Go to the SSML home page at: http://www.w3.org/TR/speech-synthesis/ and look for the link to the April 5 Working Draft.

Note that the 4SML subset is based on this version of the specification.

4SML extensions

As mentioned before, 4SML is a subset of SSML. However, 4SML can also support its own set of elements/attributes to enhance the performance of ScanSoft's TTS engines. These are the so-called 4SML extensions. They are typically expressed by the prefix 'ssft-' in front of the name of the element/attribute. At this moment the supported extensions are:

- ssft-dtype attribute of speak, paragraph and sentence with values "EMail" or "Text". With this attribute, the user can toggle the TTS output behavior between normal text mode or e-mail specific mode.

- ssft-domaintype attribute of speak, paragraph and sentence (value e.g. "propernames"). With this attribute Scansoft vendor specific domain types can be deployed.

Chapter III

To enable SSML support, set the mark up type via TtsSetParam(Ex):

```
TtsSetParam(hTtsInst, TTS_MARKUP_TYPE_PARAM ,
MARKUP_4SML);
```

By default, the mark up type is set to none.

```
TtsSetParam(hTtsInst, TTS_MARKUP_TYPE_PARAM ,
MARKUP_NONE);
```

The section below describes SSML support included in the Telecom RealSpeak host V4.0 – Brazilian Portuguese language version.

Overview of the supported 4SML tags for Brazilian Portuguese

XML encoding types for Brazilian Portuguese

The encoding is specified in the XML text declaration ("<?xml … ?>") by the encoding declaration which is of the form encoding="<EncodingName>".
E.g. <?xml version="1.0" encoding="UTF-8"?>

Telecom RealSpeak Host V4.0 – Brazilian Portuguese supports:

- "Windows-1252" and its subsets "ISO-8859-1" to "ISO-8859-9"
- The Unicode encoding "UTF-8", "UTF-16" and "UCS-4" (Note that the alias "ISO-10646-UCS-4" is not supported)

Chapter III

Elements

- paragraph or p: support for attributes “ssft-domaintype” and “ssft-dtype” (see “4SML Extensions”)
- sentence or s: support for attributes “ssft-domaintype” and “ssft-dtype” (see “4SML Extensions”)
- phoneme, with support for the unipa alphabet
- sub
- voice: support for attributes “xml:lang”, “name”, “gender” and “variant”.
- emphasis
- break
- prosody: support for rate and volume
- mark: only support for empty elements where the name attribute value is a positive number.

Chapter III

4SML Tag List for Brazilian Portuguese

For reasons of compatibility with the 'standard' Brazilian Portuguese system, the parallel text control sequence (<esc> sequence) is listed where applicable. As such, a similar TTS behavior can be created – or combined - with non-SSML driven text input.

4SML Tags	Comment	Corresponding control sequence
High-level and document structure tags		
<speak>	Supported	
xml:lang	Supported 'pt-BR' for Brazilian Portuguese. Attribute of speak, paragraph, sentence and voice. The value specifies the language in the "xml:lang" attribute (XML v1) format (<lang>-<country>).	
ssft-dtype="<doctype>"	Supported Attribute of speak, paragraph and sentence. Supported values are "EMail" and "Text" (these strings are case-sensitive).	<esc>%x
ssft-domaintype="xxx"	Not supported. To use this tag, a specific custom G2p needs to be installed. See the next section of this manual for further explanation. Attribute of speak, paragraph and sentence. value e.g. "propernames" or "Finance".	<esc>\domain=xxx\ …<esc>\domain\

Chapter III

4SML Tags	Comment	Corresponding control sequence
High-level and document structure tags		
<paragraph>	Supported <paragraph> </paragraph>	<esc>E
<p>	Supported Same as <paragraph>	<esc>E
<sentence>	Supported <sentence> </sentence>	<esc>E
<s>	Supported Same as <sentence>	<esc>E

4SML Tags	Comment	Corresponding control sequence
Text normalization tags		
<say-as interpret-as=”number” format=”cardinal” >	Some types are not supported, see below.	<esc>\tn=number_cardinal\
<say-as interpret-as=”number” format=”digits” >	Supported	<esc>\ tn=number_digits\
<say-as interpret-as=”number” format=”decimal” >	Supported	<esc>\ tn=number_decimal\
<say-as interpret-as=”number”>	Supported	<esc>\ tn=number\
<say-as interpret-as=”number” format=”ordinal” >	Supported	<esc>\ tn=number_ordinal\
<say-as interpret-as=”number” format=”telephone”>	Supported	<esc>\ tn=number_telephone\
<say-as interpret-as=”number” format=”telephone” detail= “punctuation”>	Supported	<esc>\ tn=number_telephone_punctuation\
<say-as interpret-as=”ordinal>”	Supported	<esc>\ tn=ordinal\
<say-as interpret-as=”acronym”>	Supported	<esc>\ tn=acronym\

Chapter III

4SML Tags	Comment	Corresponding control sequence
Text normalization tags		
<say-as interpret-as=”acronym” detail=”strict”>	Supported	
<say-as type=”spell-out”>	Supported	<esc>\ tn=spell\…
<say-as interpret-as=”measure”>	Not supported	
<say-as interpret-as=”letters”>	Not supported	
<say-as interpret-as=”letters” detail=”strict”>	Not supported	
<say-as interpret-as=”words”>	Not supported	
<say-as interpret-as=”date”>	Not supported	
<say-as interpret-as=”date” format=”mdy”>	Not supported	
<say-as interpret-as=”date” format =”dmy”>	Not supported	
<say-as interpret-as=”date” format=”ymd”>	Not supported	
<say-as interpret-as=”date” format=”ym”>	Not supported	
<say-as interpret-as”date” format=”my”>	Not supported	
<say-as interpret-as=”date” format=”dm”>	Not supported	
<say-as interpret-as=”date” format=”md”>	Not supported	
<say-as interpret-as=”date” format=”y”>	Not supported	
<say-as interpret-as=”date” format=”m”>	Not supported	
<say-as interpret-as=”date” format=”d”>	Not supported	
<say-as interpret-as=”time”>	Not supported	
<say-as interpret-as=”time” format=”h”>	Not supported	
<say-as interpret-as=”time” format=”hm”>	Not supported	

Chapter III

4SML Tags	Comment	Corresponding control sequence
Text normalization tags		
<say-as interpret as="time" format="hms">	Not supported	
<say-as interpret-as="duration" format="hms">	Not supported	
<say-as interpret-as="duration" format="hm">	Not supported	
<say-as interpret-as="duration" format="ms">	Not supported	
<say-as interpret-as="duration" format="h">	Not supported	
<say-as interpret-as="duration" format="m">	Not supported	
<say-as interpret-as="duration" format="s">	Not supported	
<say-as interpret-as="duration">	Not supported	
<say-as interpret-as="currency">	Not supported	
<say-as interpret-as="telephone">	Not supported	
<say-as interpret-as="telephone" detail="punctuation">	Not supported	
<say-as interpret-as="address">	Not supported	
<say-as interpret-as="spell">	Not supported	
<say-as interpret-as="name">	Not supported	
<say-as interpret-as="net" format="email">	Not supported	
<say-as interpret-as="net" format="uri">	Not supported	
<say-as interpret-as="net">	Not supported	

Chapter III

4SML Tags	Comment	Corresponding control sequence
Pronunciation tags		
<phoneme alphabet="ipa">	Not supported	
<phoneme alphabet="unipa">	Supported See section 'the Brazilian Portuguese L&H+ and UNIPA phonetic alphabets' for an overview of the alphabet.	
<sub alias="xxxxx">	Supported _{YYY} -→ xxx	

4SML Tags	Comment	Corresponding control sequence
Prosody and style tags		
<voice name="xxx">	Not supported	
<voice gender="xxx">	Not supported	
<voice variant="x">	Not supported	
<voice age="xx">	Not supported	
<voice xml:lang="xxx">	pt-BR supported (See xml:lang attribute described earlier)	
<emphasis level="default">	Supported Same as level= "moderate".	
<emphasis level="strong">	Supported	
<emphasis level="moderate">	Supported	
<emphasis level="reduced">	Supported	

Chapter III

4SML Tags	Comment	Corresponding control sequence
Prosody and style tags		
<emphasis level="none">	Supported (but no effect)	
<break size="default">	Supported Language and voice dependent Same effect as <break size="medium">	<esc>\pause=xxx\
<break size="none">	Supported (but no effect)	<esc>\pause=0\
<break size="small">	Supported Language and voice dependent	<esc>\pause=xxx\
<break size="medium">	Supported Language and voice dependent	<esc>\pause=xxx\
<break size="large">	Supported Language and voice dependent	<esc>\pause=xxx\
<break time="*xxx*ms"> (milliseconds) <break time="*x*s"> (seconds)	Supported	<esc>\pause=yyy\ (yyy is pause duration in milliseconds)
<prosody volume="*xx*">	Supported See section The Volume Scale for more details.	<esc>\vol=xx\ with xx in [0-100].
<prosody volume="+*xx*%"> <prosody volume="-*xx*%">	Supported See section "The Volume Scale" for more details.	<esc>\vol=xx\ with xx in [0-100].
<prosody volume="+*xx*"> <prosody volume="-*xx*">	Supported	<esc>\vol=x\ with xx in [0-100]

Chapter III

4SML Tags	Comment	Corresponding control sequence
Prosody and style tags		
<prosody volume="default">	Supported	<esc>\vol=80\
<prosody volume="silent">	Supported	<esc>\vol=0\
<prosody volume="soft">	Supported	<esc>\vol=40\
<prosody volume="medium">	Supported	<esc>\vol=70\
<prosody volume="loud">	Supported	<esc>\vol=95\
<prosody pitch="default">	Not supported	
<prosody pitch="high">	Not supported	
<prosody pitch="medium">	Not supported	
<prosody pitch="low">	Not supported	
<prosody pitch="xxx%">	Not supported	
<prosody pitch="xxx">	Not supported	
<prosody contour="xxxx">	Not supported	
<prosody range="xxx">	Not supported	
<prosody range="xxx%">	Not supported	
<prosody range="default">	Not supported	
<prosody range="low">	Not supported	

Chapter III

4SML Tags	Comment	Corresponding control sequence
Prosody and style tags		
<prosody range="medium">	Not supported	
<prosody range="high">	Not supported	
<prosody rate="*xxx*">	Supported	<esc>\rate_wpm=xxx\ (rate in words per minute)
<prosody rate="+*xx*%"> <prosody rate="-*xx*%">	Supported	<esc>\rate=yy\ with yy in [1-100] or <esc>\rate_wpm=yy\
<prosody rate="+*xx*"> <prosody rate="-*xx*">	Supported	<esc>\rate=yy\ with yy in [1-100] or <esc>\rate_wpm=yy\
<prosody rate="default">	Supported	<esc>\rate=50\
<prosody rate="slow">	Supported	<esc>\rate=20\
<prosody rate="medium">	Supported	<esc>\rate=50\
<prosody rate="fast">	Supported	<esc>\rate=70\
<audio>	Supported	<audio>
<prosody duration="xxx">	Not supported	

4SML Tags	Comment	Corresponding control sequence
Other tags		
<mark name="xxx">	Partially supported. Only supported are empty elements where the name attribute value is a positive number.	

Chapter III

The Volume Scale

The volume can be specified as a floating-point number in the range 0.0 to 100.00 using <prosody volume="xx">. The two other representations are: a descriptive term like "medium" and relative changes (e.g. <volume ="+15%">). The SSML spec of W3C does not fully specify how values have to be interpreted: it merely says that higher values are louder and that the value zero is equivalent to specifying "silent".

We use a logarithmic scale (like the dB scale).

4SML value	Amplification factor	Loudness in dB
100	2.00	+6dB
90	1.41	+3dB
80	1.00	+0dB
70	0.71	-3dB
60	0.50	-6dB
50	0.35	-9dB
40	0.25	-12dB
30	0.18	-15dB
20	0.125	-18dB
10	0.088	-21dB
0	0.00	-∞ dB

The formula for the dB value is:

$$x \text{ (in dB)} = 20 * \log10(\text{Amplification})$$

The formula for converting the 4SML value y to the dB value x:

$$x \text{ (dB)} = (y - 80) * 0.3 \text{ dB}$$

And the formula to derive the 4SML value from the dB value x:

$$y \text{ (4SML)} = (x / 0.3) + 80$$

The following table specifies the conversion table for the descriptive volume levels. This specification is subject to change.

Chapter III

4SML descriptive volume level	4SML numerical volume value	Amplification factor	Loudness in dB
loud	95	1.68	+4.5dB
default	80	1.00	+0dB
medium	70	0.71	-3dB
soft	40	0.20	-12dB
silent	0	0.00	-∞ dB

Telecom RealSpeak / Host SDK

Chapter IV

Custom G2P Dictionaries

User's Guide for Brazilian Portuguese
V4.0

Chapter IV

Custom G2P Dictionaries

Introduction

ScanSoft's RealSpeak system now offers support for custom G2P dictionaries. A custom G2P dictionary module is an add-on module specifically designed to improve the quality of pronunciation for specific kinds of words.

The Brazilian Portuguese system is currently not designed to support the use of a custom G2P dictionary module.

Telecom RealSpeak / Host SDK

Chapter V

Appendix

User's Guide for Brazilian Portuguese
V4.0

Chapter V

Appendix

Appendix: Brazilian Portuguese voice name

The Telecom RealSpeak /Host Text-To-Speech system now supports selecting the voice and language via a string as well as a define (please see the definition for the function **TtsInitialize()** in the `Programmers Guide` and also the `Backwards Compatibility Guide` for details). The name strings for the currently supported Brazilian Portuguese voices are listed in the table below.

Brazilian Portuguese Voice Name Strings	
Voice	**Name String**
Raquel	"Raquel"

The string to use to set the language to Brazilian Portuguese is "Brazilian Portuguese."