
Language Supplement



Speechify 3.0 for fr-FR
(French, France)

Document History

Date	Release Name
August 2003	Third Edition – for Speechify 3.0
January 2003	Second Edition – for Speechify 2.1.5
September 2002	First Edition, Update 4 – for Speechify 2.1.3
June 2002	First Edition, Update 3 – for Speechify 2.1
April 2002	First Edition, Update 2 – for Speechify 2.n
March 2002	First Edition, Update – for Speechify 2.0
January 2002	First Edition – for Speechify 2.0

Notice

Copyright © 1995–2003 by SpeechWorks International, Inc. All rights reserved.

The information in this document is subject to change without notice.

U.S. Patent Nos. 5,995,928; 5,809,494; 5,765,130; 6,061,651; and 6,173,266. One or more patents may be pending in the United States and other countries.

SpeechWorks, DialogModules, SADL, SMARTRecognizer, SpeechCare, SpeechCookie, Speechify, SpeechSecure, SpeechSite, SpeechSpot, SpeechWorks Here, the SpeechWorks logo and the SpeechWorks Here logo are trademarks or registered trademarks of SpeechWorks International, Inc. in the United States and other countries. All other trademarks are property of their respective owners. Windows NT is a registered trademark of Microsoft Corporation.

Portions of Speechify software are subject to copyrights of GlobeTrotter Software, Inc.

Published by:

SpeechWorks International, Inc.
695 Atlantic Avenue
Boston, MA 02111

Table of Contents

1. Introduction

Overview	1
New and changed information	1
Getting started	2

2. Embedded Tags

Pronouncing postal addresses.....	3
-----------------------------------	---

3. Symbolic Phonetic Representations

Vowels.....	7
Consonants	8
Syllable stress.....	9
Syllable boundary	9
Liaison	10

4. Phoneme marks

5. User Dictionaries

Main dictionary.....	13
Valid main dictionary entries.....	14
Main dictionary examples	14
Abbreviation dictionary	15
Valid abbreviation dictionary entries	15
Abbreviation dictionary examples.....	15
Root dictionary	16
Allowable root dictionary entries.....	16
Root dictionary examples	16

6. W3C SSML



Introduction

Overview

This document contains language-specific reference information for application developers using SpeechWorks Speechify 3.0 and higher.

Speechify™ is a Text-To-Speech (TTS) system.

New and changed information

Information about installing voices has been moved from this supplement to the *Speechify User's Guide*.

There are small improvements throughout this document.

See the *Speechify Migration Guide* for a complete list of changes in Speechify 3.0.

Getting started

To get started, we recommend that you are familiar with the *Speechify User's Guide*, which provides comprehensive installation, programming, and reference information about the Speechify product. You should also review the release notes distributed with this product for the latest information, restrictions, and known problems.

Other documents in the Speechify documentation set:

- ❑ *Speechify E-Mail Pre-processor Developer's Guide* explains the types of input that the SpeechWorks E-mail Pre-processor handles, how it processes the messages, the modes that the application can take advantage of at runtime, the layout and use of the E-mail substitution dictionary, and provides the API function prototypes, types, error codes and constants.
- ❑ *Speechify Migration Guide* introduces several new features that have been added to Speechify 3.0 relative to Speechify 2.x releases.
- ❑ The *SpeechWorks Licensing Handbook* describes the process for getting licenses to run the Speechify software, details about configuring your license server, and related topics.
- ❑ There is a *Speechify Language Supplement* for each supported language.

Support services

To receive technical support from SpeechWorks International, Inc., use the following methods:

- ❑ Check the FAQs, visit the Knowledge Base, or ask a question at:

<http://techsupport.speechworks.com/>

- ❑ Ask for “technical support” at +1 617 428-4444



Embedded Tags

Embedded tags are special codes that can be inserted into input text to customize Speechify's behavior in a variety of ways. The *Speechify User's Guide* describes the Speechify tag format and functionality. Below is the description of a tag that is specific to French French.

Pronouncing postal addresses

Use “address mode” to improve the interpretation of postal addresses for France. The address tags enclose a single address or a list of addresses:

Tag	Description
\addr1	Begin address mode.
\addr0	End address mode.

Address mode remains in effect until turned off with the end tag (or until the end of the speak request).

Address formats

Addresses have the following parts. *Separate each part using commas or newlines.*

- ❑ Building information – Optional. Residence or building name or number.
- ❑ Street address – A street number and name with typical abbreviations.

- ❑ Post office box number – Optional. Appears after, or in place of, a street address.
- ❑ Postal Code-City

Use address delimiters to improve the speech

To improve the resulting speech, insert newline or comma delimiters between address parts. Otherwise, any ambiguities may degrade the spoken output.

If the street address contains an apartment, suite, floor, or building number, you can improve the speech by separating the element with a newline or a comma. The following examples are pronounced in the same way:

```
bât C, résid Les Amandiers, 1 r Sommeiller, 34070 MONTPELLIER
```

```
bât C  
résid Les Amandiers  
1 r Sommeiller  
34070 MONTPELLIER
```

In a list of addresses (where address mode is turned ON for more than one address), each address should end with a double newline. Alternatively, you can enclose each address separately within address tags. That is, the following are equivalent:

```
\!addr1  
42 Av. des Tilleuls, 10120 St André les Vergers
```

```
12 Fbg de Champagne, 10260 FOUCHERES  
\!addr0
```

```
\!addr1 42 Av. des Tilleuls, 10120 St André les Vergers \!addr0  
\!addr1 12 Fbg de Champagne, 10260 FOUCHERES \!addr0
```

Examples of addresses

Comma-delimited:

```
10 Av Clemenceau, 68100 MULHOUSE  
bât C, résid Les Amandiers, 1 r Sommeiller, 34070 MONTPELLIER
```

Newline-delimited:

```
5 r Gén Blaise  
75011 Paris
```


141 Bis av de Limoges
87270, COUZEIX

27 Av. de plantières BP 10
57998 METZ-ARMEES

bât K8
8, r Honoré de Balzac
62100 Calais

Avoid extraneous text

Once address mode begins, all text is assumed to be part of the address, so you should avoid including non-address text inside the tags. An example:

```
\!addr1 Ce thé sent le jasmin. \!addr0
```

The string above is read “Ce thé sentier le jasmin.”

However, addresses may include a limited amount of additional text such as company or building names. For example:

Cabinet médical des Bordières
6 place des Bouleaux
94000 CRETEIL

How numbers are processed

The following rules apply to cardinal numbers in addresses, such as house numbers, building numbers, and postal codes. Ordinal numbers within addresses are read as usual (for example, pl Albert 1er is read “place Albert premier”).

- ❑ 5 digits – If the digits end in zeros (xy000), the number is read “xy mille.” For example, 18000 is read “dix-huit mille.”

Otherwise, the number is grouped into “vw, xyz.” (For example, 25435 is read “vingt-cinq, quatre cents trente-cinq.”) Leading zeros in either group are read as “zéro.” (For example, 02130 is read “zéro deux, cent trente,” and 75013 is read “soixante-quinze, zéro treize.”)

- ❑ Other numbers – Read as ordinary numbers.

Prosody

Address mode inserts phrase boundaries and accompanying pauses as follows:

- ❑ Between the number and the street within the street address.
- ❑ Between the street address and the post office box, if any.
- ❑ Between the street address or post office box and the postal code.
- ❑ After the first two numbers of a postal code (unless ending in 000, as in 18000).
- ❑ Between the postal code and the city.
- ❑ Wherever else a comma appears in the address.

For example:

```
\!addr1 Bât A, 13 all Fernand Léger BP 110, 75014 Paris. \!addr0
```

Above, the address is read as follows (pauses are indicated with the commas):

```
bâtiment A, treize, Allée Fernand Léger, B P cent dix,  
soixante-quinze, zéro quatorze, Paris.
```

Address mode also inserts a sentence boundary at the end of each address when the address is delimited by a double newline. No sentence boundaries are inserted within an address.

Address fragments

Addresses do not need to be complete.

The input text can contain address fragments (for example, either the street address part or the postal code-city part).



Symbolic Phonetic Representations

The following tables show the inventory of available symbols for use in fr-FR Symbolic Phonetic Representations (SPRs). Use this chapter in conjunction with the SPR information found in the *Speechify User's Guide*.

Each sound symbol is accompanied by examples illustrating typical spellings of the sound in actual words, with the letters representing the given sound bolded. Due to dialectal differences, the SPR examples shown may not always agree with your own pronunciations.

Vowels

The following table includes the fr-FR symbols for vowels.

French symbol	Example words
i	f ilm, t ype, r ythmique
e	café, dé f ormer, é t é
E	père, annua i re, m e r
y	u tile, p ure, B runo
'eu'	p eu, j eu n er, émeute
'oe'	p eu r , j eu n e
u	r oue, o ù, a oût, t our
o	p aule, t ôt, e aux
c	p aul, n ote, échalotte

French symbol	Example words
a	p attes, l ac, c ave
x	litres, marbre
'E~'	f in, pl ein, f aim
'o~'	bon , p ont
'a~'	banc , en , temps

Consonants

The following table includes the fr-FR symbols for consonants.

French symbol	Example words
p	p orte, ap porter
b	b ébé
t	ton , pat te, th éâtre
d	d ort, add ition
k	k ilo, cal er, qu ai
g	gu erre, bag ue, gar er
f	ch ef, f aim, ph are
v	lav er, w agon, vis iter
s	s ans, amb ition, fa çon
z	j aser, zig zaguer
S	ch eval, sch éma
Z	rag e, g îte, j ouer
m	m aman, f emme
n	An ne, an onyme
'nj'	ag neau, camp agne
'ng'	park ing
l	lit re, ill isible, p âle

French symbol	Example words
r	r are, car r eau, chèv r e
j	hiérarchie, paille, y éyé
H	nuit, nu é e
w	o ui, vo l à

Syllable stress

You can mark syllables for stress with a digit. The following table includes the fr-FR symbols for syllable stress.

0	no stress
1	primary stress (most prominent stress in the word)
2	secondary stress

If a word has more than one syllable, you can mark one of these syllables for primary stress, and mark others for secondary stress or no stress.

A syllable that is not marked for stress is assumed to have no stress, unless it is the only syllable of a word, in which case it is assigned a primary stress. In French, you must place the stress marker immediately to the left of the vowel of the syllable.

Syllable boundary

The following table includes the fr-FR symbol for a syllable boundary.

.	Beginning of a syllable
---	-------------------------

In French, a single syllable can cross a word boundaries. The syllabification of the sequence is shown below; a dot (.) indicates a syllable boundary and the pound symbol (#) indicates a word boundary.

E . m#E . m#Em

Liaison

An underscore can be used following a word-final consonant (but within the right bracket which closes the SPR) to indicate that it is a liaison consonant. The consonant is pronounced if and only if there is a following word in the same phrase beginning with a vowel.

For example, a root dictionary key *petit* with the translation value `\![p'oe't1it_]` has the final [t] pronounced in the input string *un petit ami* but not in the input string *un petit chien*. An entry *net* with the translation value `\![nEt]` has the final [t] pronounced regardless of context.

The following examples show how to use the symbol for liaison.

<code>_</code>	Allow liaison if the following word begins with a vowel. For example:
<code>\![p0'oe't1it_]</code>	The [t] is pronounced only when the following word begins with a vowel.
<code>\![nEt]</code>	The [t] is always pronounced.



Phoneme marks

The phoneme-mark data structure is useful for matching each moment of synthesized speech with the phonemes being spoken. This information is particularly useful when you synchronize facial animation with speech output.

The structure is returned to your application's callback function, which is described under “SWIttsCallback()” in the API reference section of the *Speechify User's Guide*. Each phoneme-mark contains the phoneme name, along with other information about the phoneme. The phoneme-mark symbols for fr-FR are documented in the following table.

Note: The phoneme names used in phoneme marks are not necessarily the same as those used for SPR input.

Phone	Examples
a	p attes, lac , cave
a~	ban c, en
b	b alle, ro be
c	pau l, no te
d	d ort, add ition
e	café, passer
E	mer , annua ir e
eu	peu , jeû ner
E~	fai m, plein
f	chef , fai m

Phone	Examples
g	g uerre, bag ue
H	lui , nuit
i	film , rythmiqu e
j	paille , yoga
k	kilo , lac
l	litre , illisible
m	maman , femme
n	Anne , ni
ng	park ing , camp ing
nj	ag ne au, campag ne
o	paule , eaux
oe	peur , jeune
o~	bon , pont
p	porte , soupe
pau	[Indicates a pause]
r	parer , rare
s	sans , ambition
S	cheval , acheter
t	ton , patte
u	roue , tour
v	laver , visiter
w	oui , moi
x	litres , marbre
y	utile , pure
z	jaser , zigzag
Z	rage , jouer



User Dictionaries

This chapter describes French French behaviors for Speechify dictionaries. Use this chapter in conjunction with the dictionary information found in the *Speechify User's Guide*.

Lookups in the French French user dictionary ignore cliticized articles, object pronouns, and reflexive pronouns. For example, if *apple* is an entry in any of the user dictionaries, it matches an input string such as *l'apple*.

The maximum length for French French dictionary keys is 128 characters. The maximum translation length is 512 characters.

Main dictionary

The main dictionary is an all-purpose user dictionary for replacing a word in an input text with almost any type of input string. Main dictionary entries are case-sensitive.

Valid main dictionary entries

The following table summarizes the valid main dictionary keys and translations:

Key	Translation
Latin 1 letters, both upper and lower case, digits. Non-alphanumeric characters, including: @, #, \$, %, &, *, + Apostrophes, quotation marks, parentheses, brackets, etc. Punctuation, except as the final character.	Anything that is valid input to the text-to-speech engine, including whitespace, punctuation, SPRs, and embedded tags.
Disallowed: white space	Disallowed: SAPI tags, SSML tags, and bookmarks

Main dictionary examples

The following table shows examples of main dictionary entries:

Key	Translation
ab@nulle.org	a b à nulle point o r g
7/7	sept sur sept
3615	trente-six quinze
CAF	‘[s0e0a1Ef]
HEC	‘[aS0’oe’s1e]
Pin’s	‘[p1inz]
CAPES	capesse

Abbreviation dictionary

The abbreviation dictionary handles word abbreviations that translate to one or more words in ordinary spelling. The entries are case-sensitive.

Valid abbreviation dictionary entries

The following table summarizes the format of valid keys and translations in the abbreviation dictionary:

Key	Translation
Sequences of one or more letters optionally separated by periods (x.x.x or xx.xx.xx).	One or more valid words in ordinary spelling, including both upper and lower case letters, separated by white space or hyphen.
Sequences of letters, with or without the trailing period that may be considered part of the abbreviation (xxx. or xxx).	
Internal hyphen (not the first or last character in the sequence)	
Upper or lower case letters.	
Disallowed: non-alphabetic characters, white space, or punctuation (except periods).	Disallowed: digits, punctuation, SPRs, tags.

Abbreviation dictionary examples

The following table shows examples of abbreviation dictionary entries:

Key	Translation
ndlr	note de la rédaction
B.P.	boîte postale
P-D.G.	président-directeur général

Root dictionary

The root dictionary is used for ordinary words, like nouns (including proper names), verbs, or adjectives. Unlike the main and abbreviation dictionaries, it is not case-sensitive.

The fr-FR root dictionary requires that all forms of a word must be entered separately. For example, both singular and plural forms of a noun must be entered in order to produce the desired pronunciation in both words.

In French French, a root dictionary translation applies when the key occurs as the root of a larger word. For example, if *apple* is a root dictionary key, the translation value is substituted for *apple* when it occurs by itself, or as the root of words like *l'apple* and *apples*. There is no need to enter these words separately in the user root dictionary.

Allowable root dictionary entries

The following table summarizes the format of valid keys and translations in the root dictionary:

Key	Translation
A single word in ordinary spelling, all lowercase letters.	A single word in ordinary spelling. A valid SPR.
Disallowed: digits, punctuation, or other non-letter characters; white space	Disallowed: digits, punctuation, or other non-letter characters, white space, tags, or annotations

Root dictionary examples

The following table shows examples of root dictionary entries:

Key	Translation	Would apply to...
véro	Véronique	Véro, véro

Key	Translation	Would apply to...
cambridge	'[k0'a~'br1idZ]	Cambridge, cambridge
basket	'[b0ask1Et]	basket, baskets



W3C SSML

Speechify includes support for input text formatted according to the W3C Speech Synthesis Markup Language (SSML).

In order to invoke the correct processing for French French text, the given W3C SSML input must have the `xml:lang` attribute set to “fr-FR”. If this attribute is absent, Speechify assumes “en-US” as a default.

In addition to producing French French output, the W3C SSML processor for fr-FR parses and expands tag content in a manner consistent with that locale. For example, the tag `<say-as type = "currency">10,20</say-as>` is pronounced “dix euros vingt.”

For more description of W3C SSML support in Speechify, see the *Speechify User's Guide*.

Index

Symbols

\addr0 [3](#)

\addr1 [3](#)

A

abbreviation dictionary [15](#)

address mode [3](#)

annotations

 pronouncing postal addresses [3](#)

D

dictionary

 root [16](#)

E

embedded tags [3](#)

F

French SPRs [7](#), [10](#)

P

postal addresses

 annotations [3](#)

R

root dictionary [16](#)

S

SPR symbols

 French French [7](#)

support services [2](#)

