

Package ‘ipa’

July 22, 2025

Title Convert Between Phonetic Alphabets

Version 0.1.0

Description Converts character vectors between phonetic representations. Supports IPA (International Phonetic Alphabet), X-SAMPA (Extended Speech Assessment Methods Phonetic Alphabet), and ARPABET (used by the CMU Pronouncing Dictionary).

License MIT + file LICENSE

URL <https://github.com/rossellhayes/ipa>

BugReports <https://github.com/rossellhayes/ipa/issues>

Depends R (>= 2.10)

Imports stringi

Suggests covr, testthat

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

NeedsCompilation no

Author Alexander Rossell Hayes [aut, cre, cph] (ORCID:
<<https://orcid.org/0000-0001-9412-0457>>)

Maintainer Alexander Rossell Hayes <alexander@rossellhayes.com>

Repository CRAN

Date/Publication 2020-06-04 10:20:02 UTC

Contents

arpa	2
convert_phonetics	3
ipa	4
sampa	5
Index	6

`arpa`*Convert phonetic representation from ARPABET (CMU)*

Description

Converts to IPA or X-SAMPA

Usage

```
arpa(x, to = c("ipa", "xsampa"))
```

```
arpabet(x, to = c("ipa", "xsampa"))
```

```
cmu(x, to = c("ipa", "xsampa"))
```

Arguments

<code>x</code>	A phonetic character vector in ARPABET
<code>to</code>	The phonetic alphabet to convert to. Defaults to "ipa". "sampa" is an accepted alias for "xsampa".

Value

A character vector the same length as `x`

See Also

[convert_phonetics\(\)](#) to convert between any phonetic representations

[ipa\(\)](#) to convert from IPA

[sampa\(\)](#) and [xsampa\(\)](#) to convert from X-SAMPA

Examples

```
arpa('HH EH L OW')
arpa(c('HH EH L OW', 'W ER L D'))

arpa('HH EH L OW', "xsampa")
arpa(c('HH EH L OW', 'W ER L D'), "xsampa")
```

convert_phonetics	<i>Convert between phonetic representations</i>
-------------------	---

Description

Converts between IPA, X-SAMPA and ARPABET (used by the CMU Pronouncing Dictionary)

Usage

```
convert_phonetics(
  x,
  from = c("ipa", "xsampa", "arpabet"),
  to = c("ipa", "xsampa", "arpabet")
)
```

Arguments

x	A phonetic character vector in IPA, X-SAMPA, or ARPABET
from	The phonetic alphabet to convert from
to	The phonetic alphabet to convert to. "sampa" is an accepted alias for "xsampa". "arpa" and "cmu" are accepted aliases for "arpabet".

Value

A character vector the same length as x

See Also

[ipa\(\)](#) to convert from IPA

[sampa\(\)](#) and [xsampa\(\)](#) to convert from X-SAMPA

[arpa\(\)](#), [arpabet\(\)](#) and [cmu\(\)](#) to convert from ARPABET (CMU)

Examples

```
convert_phonetics('%hE"loU', from = "xsampa", to = "ipa")
convert_phonetics(c('%hE"loU', "w3`ld"), from = "xsampa", to = "ipa")
convert_phonetics('%hE"loU', from = "xsampa", to = "arpabet")
convert_phonetics(c('%hE"loU', "w3`ld"), from = "xsampa", to = "arpabet")

convert_phonetics('HH EH L OW', from = "arpabet", to = "ipa")
convert_phonetics(c('HH EH L OW', "W ER L D"), from = "arpabet", to = "ipa")
convert_phonetics('HH EH L OW', from = "arpabet", to = "xsampa")
convert_phonetics(c('HH EH L OW', "W ER L D"), from = "arpabet", to = "xsampa")

# `convert_phonetics()` supports raw Unicode input or escaped Unicode sequences
# (\uxxxx), but raw Unicode is not supported in R package documentation,
# hence the ugly examples
```

```

convert_phonetics("\u02cch\u025b\u028a", from = "ipa", to = "xsampa")
convert_phonetics(
  c("\u02cch\u025b\u028a", "w\u025dld"), from = "ipa", to = "xsampa"
)
convert_phonetics("\u02cch\u025b\u028a", from = "ipa", to = "arpabet")
convert_phonetics(
  c("\u02cch\u025b\u028a", "w\u025dld"), from = "ipa", to = "arpabet"
)

```

ipa

*Convert phonetic representation from IPA***Description**

Converts to X-SAMPA or ARPABET (used by the CMU Pronouncing Dictionary)

Usage

```
ipa(x, to = c("xsampa", "arpabet"))
```

Arguments

x	A phonetic character vector in IPA
to	The phonetic alphabet to convert to. Defaults to "xsampa". "sampa" is an accepted alias for "xsampa". "arpa" and "cmu" are accepted aliases for "arpabet".

Value

A character vector the same length as x

See Also

[convert_phonetics\(\)](#) to convert between any phonetic representations

[sampa\(\)](#) and [xsampa\(\)](#) to convert from X-SAMPA

[arpa\(\)](#), [arpabet\(\)](#) and [cmu\(\)](#) to convert from ARPABET (CMU)

Examples

```

# `ipa()` supports raw Unicode input or escaped Unicode sequences (\uxxxx),
# but raw Unicode is not supported in R package documentation,
# hence the ugly examples
ipa("\u02cch\u025b\u028a")
ipa(c("\u02cch\u025b\u028a", "w\u025dld"))

ipa("\u02cch\u025b\u028a", to = "arpabet")
ipa(c("\u02cch\u025b\u028a", "w\u025dld"), to = "arpabet")

```

sampa	<i>Convert phonetic representation from X-SAMPA</i>
-------	---

Description

Converts to IPA or ARPABET (used by the CMU Pronouncing Dictionary)

Usage

```
sampa(x, to = c("ipa", "arpabet"))
```

```
xsampa(x, to = c("ipa", "arpabet"))
```

Arguments

x	A phonetic character vector in X-SAMPA
to	The phonetic alphabet to convert to. Defaults to "ipa". "arpa" and "cmu" are accepted aliases for "arpabet".

Value

A character vector the same length as x

See Also

[convert_phonetics\(\)](#) to convert between any phonetic representations

[ipa\(\)](#) to convert from IPA

[arpa\(\)](#), [arpabet\(\)](#) and [cmu\(\)](#) to convert from ARPABET (CMU)

Examples

```
sampa('%hE"loU')
sampa(c('%hE"loU', "w3`ld"))

sampa('%hE"loU', "arpabet")
sampa(c('%hE"loU', "w3`ld"), "arpabet")
```

Index

arpa, [2](#)
arpa(), [3–5](#)
arpabet (arpa), [2](#)
arpabet(), [3–5](#)

cmu (arpa), [2](#)
cmu(), [3–5](#)
convert_phonetics, [3](#)
convert_phonetics(), [2, 4, 5](#)

ipa, [4](#)
ipa(), [2, 3, 5](#)

sampa, [5](#)
sampa(), [2–4](#)

xsampa (sampa), [5](#)
xsampa(), [2–4](#)