Package ‘childesr’

October 17, 2019

Type Package

Title Accessing the ‘CHILDES’ Database

Description Tools for connecting to ‘CHILDES’, an open repository for transcripts of parent-child interaction. For more information on the underlying data, see <http://childes-db.stanford.edu>.

Version 0.1.2

Depends R (>= 3.4.0)

License GPL-3

URL https://github.com/langcog/childesr

BugReports http://github.com/langcog/childesr/issues

Encoding UTF-8

LazyData true

Imports DBI (>= 0.8), dbplyr (>= 1.2.1), dplyr (>= 0.7.4), jsonlite (>= 1.5), magrittr (>= 1.5), purrr (>= 0.2.4), RMySQL (>= 0.10.14)

Suggests knitr, rmarkdown, curl

RoxygenNote 6.1.1

VignetteBuilder knitr

NeedsCompilation no

Author Mika Braginsky [aut, cre], Alessandro Sanchez [aut, ctb], Daniel Yurovsky [aut], Kyle MacDonald [ctb], Stephan Meylan [ctb]

Maintainer Mika Braginsky <mika.br@gmail.com>

Repository CRAN

Date/Publication 2019-10-17 04:50:02 UTC
check_connection

Description
Check if connecting to childes db is possible

Usage
check_connection(db_version = "current", db_args = NULL)

Arguments
- db_version: String of the name of database version to use
- db_args: List with host, user, and password defined

Value
Logical indicating whether a connection was successfully formed
clear_connections

Clear all MySQL connections

description

Clear all MySQL connections

usage

clear_connections()

connect_to_childes

Connect to CHILDES

description

Connect to CHILDES

usage

connect_to_childes(db_version = "current", db_args = NULL)

arguments

db_version String of the name of database version to use
db_args List with host, user, and password defined

value

con A DBIConnection object for the CHILDES database

examples

con <- connect_to_childes(db_version = "current", db_args = NULL)
DBI::dbDisconnect(con)
get_collections

Description
Get collections

Usage

get_collections(connection = NULL, db_version = "current",
db_args = NULL)

Arguments

connection A connection to the CHILDES database
db_version String of the name of database version to use
db_args List with host, user, and password defined

Value
A ‘tbl’ of Collection data. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

Examples

get_collections()

get_content

Description
Get content

Usage

get_content(content_type, collection = NULL, language = NULL,
corpus = NULL, role = NULL, role_exclude = NULL, age = NULL,
sex = NULL, target_child = NULL, token = NULL, stem = NULL,
part_of_speech = NULL, connection)
**get_contexts**

*Get the utterances surrounding a token(s)*

**Arguments**

- **content_type**: One of "token" or "utterance"
- **collection**: A character vector of one or more names of collections
- **language**: A character vector of one or more languages
- **corpus**: A character vector of one or more names of corpora
- **role**: A character vector of one or more roles to include
- **role_exclude**: A character vector of one or more roles to exclude
- **age**: A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
- **sex**: A character vector of values "male" and/or "female"
- **target_child**: A character vector of one or more names of children
- **token**: A character vector of one or more token patterns (‘%’ matches any number of wildcard characters, ‘_’ matches exactly one wildcard character)
- **stem**: A character vector of one or more stems
- **part_of_speech**: A character vector of one or more parts of speech
- **connection**: A connection to the CHILDES database

**Description**

Get the utterances surrounding a token(s)

**Usage**

```r
get_contexts(collection = NULL, language = NULL, corpus = NULL, role = NULL, role_exclude = NULL, age = NULL, sex = NULL, target_child = NULL, token, window = c(0, 0), remove_duplicates = TRUE, connection = NULL, db_version = "current", db_args = NULL)
```

**Arguments**

- **collection**: A character vector of one or more names of collections
- **language**: A character vector of one or more languages
- **corpus**: A character vector of one or more names of corpora
- **role**: A character vector of one or more roles to include
- **role_exclude**: A character vector of one or more roles to exclude
- **age**: A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
sex: A character vector of values "male" and/or "female"

target_child: A character vector of one or more names of children

token: A character vector of one or more token patterns ('%' matches any number of wildcard characters, '_' matches exactly one wildcard character)

window: A length 2 numeric vector of how many utterances before and after each utterance containing the target token to retrieve

remove_duplicates: A boolean indicating whether to remove duplicate utterances from the results

connection: A connection to the CHILDES database

db_version: String of the name of database version to use

db_args: List with host, user, and password defined

Value

A 'tbl' of Corpus data. If 'connection' is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

Examples

```
get_contexts(target_child = "Shem", token = "dog")
```

Description

Get corpora

Usage

```
get_corpora(connection = NULL, db_version = "current",
            db_args = NULL)
```

Arguments

- connection: A connection to the CHILDES database
- db_version: String of the name of database version to use
- db_args: List with host, user, and password defined

Value

A ‘tbl’ of Corpus data. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.
**get_database_version**

**Examples**

```r
get_corpora()
```

---

**get_database_version**  
*Get database version*

**Description**

Get database version

**Usage**

```r
get_database_version(connection = NULL, db_version = "current", db_args = NULL)
```

**Arguments**

- `connection`: A connection to the CHILDES database
- `db_version`: String of the name of database version to use
- `db_args`: List with host, user, and password defined

**Value**

The database version as a string

**Examples**

```r
get_database_version()
```

---

**get_db_info**  
*Get information on database connection options*

**Description**

Get information on database connection options

**Usage**

```r
get_db_info()
```
get_participants

Value

List of database info: host name, current version, supported versions, historical versions, username, password

Examples

get_db_info()

get_participants collection = NULL, corpus = NULL, target_child = NULL, role = NULL, role_exclude = NULL, age = NULL, sex = NULL, connection = NULL, db_version = "current", db_args = NULL

Arguments

collection A character vector of one or more names of collections
corpus A character vector of one or more names of corpora
target_child A character vector of one or more names of children
role A character vector of one or more roles to include
role_exclude A character vector of one or more roles to exclude
age A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
sex A character vector of values "male" and/or "female"
connection A connection to the CHILDES database
db_version String of the name of database version to use
db_args List with host, user, and password defined

Value

A ‘tbl’ of Participant data, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.
get_speaker_statistics

Examples

get_participants()

tbl

get_speaker_statistics

Get speaker statistics

Description

Get speaker statistics

Usage

get_speaker_statistics(collection = NULL, corpus = NULL, target_child = NULL, role = NULL, role_exclude = NULL, age = NULL, sex = NULL, connection = NULL, db_version = "current", db_args = NULL)

Arguments

collection A character vector of one or more names of collections
corpus A character vector of one or more names of corpora
target_child A character vector of one or more names of children
role A character vector of one or more roles to include
role_exclude A character vector of one or more roles to exclude
age A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
sex A character vector of values "male" and/or "female"
connection A connection to the CHILDES database
db_version String of the name of database version to use
db_args List with host, user, and password defined

Value

A ‘tbl’ of Speaker statistics, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

Examples

get_speaker_statistics()
get_table

Description
Get table

Usage
get_table(connection, name)

Arguments
connection A connection to the CHILDES database
name String of a table name

Value
A 'tbl'

get_tokens

Description
Get tokens

Usage
get_tokens(collection = NULL, language = NULL, corpus = NULL,
target_child = NULL, role = NULL, role_exclude = NULL,
age = NULL, sex = NULL, token, stem = NULL,
part_of_speech = NULL, replace = TRUE, connection = NULL,
db_version = "current", db_args = NULL)

Arguments
collection A character vector of one or more names of collections
language A character vector of one or more languages
corpus A character vector of one or more names of corpora
target_child A character vector of one or more names of children
role A character vector of one or more roles to include
role_exclude A character vector of one or more roles to exclude
**get_transcripts**

- **age**: A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
- **sex**: A character vector of values "male" and/or "female"
- **token**: A character vector of one or more token patterns (‘%’ matches any number of wildcard characters, ‘_’ matches exactly one wildcard character)
- **stem**: A character vector of one or more stems
- **part_of_speech**: A character vector of one or more parts of speech
- **replace**: A boolean indicating whether to replace "gloss" with "replacement" (i.e. phonologically assimilated form), when available (defaults to TRUE)
- **connection**: A connection to the CHILDES database
- **db_version**: String of the name of database version to use
- **db_args**: List with host, user, and password defined

**Value**

A ‘tbl’ of Token data, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

**Examples**

```r
get_tokens(token = "dog")
```

**get_transcripts**  
*Get transcripts*

**Description**

Get transcripts

**Usage**

```r
get_transcripts(collection = NULL, corpus = NULL, target_child = NULL, connection = NULL, db_version = "current", db_args = NULL)
```

**Arguments**

- **collection**: A character vector of one or more names of collections
- **corpus**: A character vector of one or more names of corpora
- **target_child**: A character vector of one or more names of children
- **connection**: A connection to the CHILDES database
- **db_version**: String of the name of database version to use
- **db_args**: List with host, user, and password defined
**get_types**

**Value**

A ‘tbl’ of Transcript data, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

**Examples**

```r
get_transcripts()
```

---

**get_types**

**Get types**

**Description**

Get types

**Usage**

```r
get_types(collection = NULL, language = NULL, corpus = NULL, role = NULL, role_exclude = NULL, age = NULL, sex = NULL, target_child = NULL, type = NULL, connection = NULL, db_version = "current", db_args = NULL)
```

**Arguments**

- `collection`: A character vector of one or more names of collections
- `language`: A character vector of one or more languages
- `corpus`: A character vector of one or more names of corpora
- `role`: A character vector of one or more roles to include
- `role_exclude`: A character vector of one or more roles to exclude
- `age`: A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
- `sex`: A character vector of values "male" and/or "female"
- `target_child`: A character vector of one or more names of children
- `type`: A character vector of one or more type patterns (‘ number of wildcard characters, ‘_’ matches exactly one wildcard character)
- `connection`: A connection to the CHILDES database
- `db_version`: String of the name of database version to use
- `db_args`: List with host, user, and password defined

**Value**

A ‘tbl’ of Type data, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.
get_utterances

Examples

get_types()

get_utterances

Description

Get utterances

Usage

get_utterances(collection = NULL, language = NULL, corpus = NULL, role = NULL, role_exclude = NULL, age = NULL, sex = NULL, target_child = NULL, connection = NULL, db_version = "current", db_args = NULL)

Arguments

collection  A character vector of one or more names of collections
language     A character vector of one or more languages
corpus       A character vector of one or more names of corpora
role         A character vector of one or more roles to include
role_exclude A character vector of one or more roles to exclude
age          A numeric vector of an age or a min age (inclusive) and max age (exclusive) in months
sex          A character vector of values "male" and/or "female"
target_child A character vector of one or more names of children
connection   A connection to the CHILDES database
db_version   String of the name of database version to use
db_args      List with host, user, and password defined

Value

A ‘tbl’ of Utterance data, filtered down by supplied arguments. If ‘connection’ is supplied, the result remains a remote query, otherwise it is retrieved into a local tibble.

Examples

get_utterances(target_child = "Shem")
Index

check_connection, 2
clear_connections, 3
connect_to_childes, 3

get_collections, 4
get_content, 4
get_contexts, 5
get_corpora, 6
get_database_version, 7
get_db_info, 7
get_participants, 8
get_speaker_statistics, 9
get_table, 10
get_tokens, 10
get_transcripts, 11
get_types, 12
get_utterances, 13