

# Package ‘chilDesr’

November 13, 2020

**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 <<https://chilDesr-db.stanford.edu>>.

**Version** 0.2.1

**Depends** R (>= 3.6)

**License** GPL-3

**URL** <https://github.com/langcog/chilDesr>

**BugReports** <https://github.com/langcog/chilDesr/issues>

**Encoding** UTF-8

**LazyData** true

**Imports** DBI (>= 1.1.0), dbplyr (>= 1.4), dplyr (>= 1.0), jsonlite (>= 1.5), magrittr (>= 1.5), purrr (>= 0.3), RMySQL (>= 0.10.20)

**Suggests** knitr, rmarkdown, curl

**RoxygenNote** 7.1.1

**VignetteBuilder** knitr

**NeedsCompilation** no

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

**Maintainer** Mika Braginsky <[mika.br@gmail.com](mailto:mika.br@gmail.com)>

**Repository** CRAN

**Date/Publication** 2020-11-13 20:40:02 UTC

**R topics documented:**

check_connection . . . . .	2
clear_connections . . . . .	3
connect_to_childes . . . . .	3
get_collections . . . . .	4
get_content . . . . .	4
get_contexts . . . . .	5
get_corpora . . . . .	7
get_db_info . . . . .	7
get_participants . . . . .	8
get_speaker_statistics . . . . .	9
get_sql_query . . . . .	10
get_table . . . . .	11
get_tokens . . . . .	11
get_transcripts . . . . .	12
get_types . . . . .	13
get_utterances . . . . .	15
<b>Index</b>	<b>17</b>

---

check_connection	<i>Check if connecting to childes db is possible</i>
------------------	--

---

**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	<i>Get collections</i>
-----------------	------------------------

---

**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	<i>Get content</i>
-------------	--------------------

---

**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
)

```

### 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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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

---

get\_contexts

*Get the utterances surrounding a token(s)*

---

### Description

Get the utterances surrounding a token(s)

### Usage

```

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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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 Utterance data, filtered down by supplied arguments.

### Examples

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

---

`get_corpora`*Get corpora*

---

**Description**

Get corpora

**Usage**

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

**Arguments**

<code>connection</code>	A connection to the CHILDES database
<code>db_version</code>	String of the name of database version to use
<code>db_args</code>	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_corpora()
```

---

`get_db_info`*Get information on database connection options*

---

**Description**

Get information on database connection options

**Usage**

```
get_db_info()
```

**Value**

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

**Examples**

```
get_db_info()
```

---

```
get_participants      Get participants
```

---

**Description**

Get participants

**Usage**

```
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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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.

**Examples**

```
get_participants()
```

---

```
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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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_sql_query	<i>Run a SQL Query script on the CHILDES database</i>
---------------	---

---

**Description**

Run a SQL Query script on the CHILDES database

**Usage**

```
get_sql_query(
  sql_query_string,
  connection = NULL,
  db_version = "current",
  db_args = NULL
)
```

**Arguments**

sql_query_string	A valid sql query string 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**

The database after calling the supplied SQL query

**Examples**

```
get_sql_query("SELECT * FROM collection")
```

---

get_table	<i>Get table</i>
-----------	------------------

---

**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	<i>Get tokens</i>
------------	-------------------

---

**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
age	A numeric vector of an single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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**

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

---

```
get_transcripts
```

```
Get transcripts
```

---

**Description**

Get transcripts

**Usage**

```
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

**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**

```
get_transcripts()
```

---

get_types	<i>Get types</i>
-----------	------------------

---

**Description**

Get types

**Usage**

```
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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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.

### Examples

```
get_types()
```

---

get_utterances	<i>Get utterances</i>
----------------	-----------------------

---

### 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 single age value or a min age value and max age value (inclusive) in months. For a single age value, participants are returned for which that age is within their age range; for two ages, participants are returned for whose age overlaps with the interval between those two ages.
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, 7  
get\_db\_info, 7  
get\_participants, 8  
get\_speaker\_statistics, 9  
get\_sql\_query, 10  
get\_table, 11  
get\_tokens, 11  
get\_transcripts, 12  
get\_types, 13  
get\_utterances, 15