WWW.CDI-CLEX.ORG
A cross-linguistic lexical norms database
Jørgensen, Rune Nørgaard; Dale, Philip; Bleses, Dorthe; Fenson, Larry

Publication date:
2011

Document version
Early version, also known as pre-print

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 01. dec., 2018
The CLEX project has been carried out in collaboration by The CDI Advisory Board & CLEX stands for Cross Linguistic Lexical Norms, and is an extension of the LEX database. WWW.CDI-CLEX.ORG

In general, when you start looking for a word you wish to search the database, you have three immediate headers from which to begin. You can search the vocabulary section of the CDI (part of the CDI W & G + W&S) under the Query Vocabulary header; you can search the glossary section of the CDI (part of the CDI W & G) under the Query Glossary header; or you can search the other child language information from the CDI reports (part of the CDI W & G) under the Query CDI Other header.

### What is CLEX

#### Basic Principles

The database has been placed on a web server and there has been created a number of scripts so that it is possible to search from them in a web browser. The default language in CLEX is English (American) and the different CDI studies that are included are organized so that they refer to the original American CDI study, meaning that when a new language is included in CLEX, each word is given an English translation and, if possible, a link to the matching word in the American CDI form.

#### Results & Wives

- **Basic Principles:**
  - The default language in CLEX is English (American) and the different CDI studies that are included are organized so that they refer to the original American CDI study, meaning that when a new language is included in CLEX, each word is given an English translation and, if possible, a link to the matching word in the American CDI form.

- **Tables & flexibilities:**
  - Interesting data tables is straightforward. In the example below, you can see that 10 months of age, 10 children are included in the study 64.2% of these children did not understand the word cat, and 2.9% had said the word 'slæde' (sledge) is in the Danish CDI but not in the American CDI.

#### Tables & flexibilities

Table 1 illustrates the various datasets from different CDI studies in various languages. The system is created so that new languages can be included fairly easily, with the goal of facilitating cross-linguistic comparisons.

### Results & Wives

#### What is shown

A central issue when working on online dynamic data in a system like CLEX is to keep track of what is actually being shown. First step for the user is to identify the dataset that they are searching in. The first tab in each dataset is therefore dedicated to general information about the dataset. When a table has been generated all relevant information about the specific table is listed below the result, such as age span, number of items, etc, but also the specific item is listed, to minimize the risk of misunderstandings.

#### Search and selection

#### Searching in CLEX

You start your search by choosing a specific dataset. That could be the Danish infant (cross-sectional), the American toddler (cross-sectional), etc. When you have chosen a dataset you can choose between different types of search functions: single word, single word list, vocabulary, word class and norms (see below for description below).

In each function you start by picking the specific items you are interested in, and a table will be generated. However you can refine the criteria for the result, by specifying gender, age, etc. For most results it is possible to see the result as a graph as well.

Having the result form your search in a given dataset now gives you the opportunity to save the result of a search with the same criteria in another language. For the given example of kinder, it is best to see the equivalent word in Danish 'børn' and in Swedish 'barn'.

Finally you can make a direct comparison either in a graph or a table depending on which function you are using.

#### Norms

This option provides the overall normative information on vocabulary totals by age, in comparison to the tables and figures in the Technical Manual for the assessments.

#### Cross-linguistic advantage

Contrary to the functionality of the CDI website is the ability to cross-linguistically compare child language. In the CDI Vocabulary section of the website, you will find automatic mapping options under the Single Word, the Single Word List, the Vocabulary Subscale, and Item Comparison sections.

Before you begin, note that the manual cross-lingual mappings are more reliable in that they were made by an actual researcher. Semi-automated cross-lingual mappings are intrinsically less reliable because they are created by an automated comparison of the words, but have not been checked by an actual person.

In the example to the above, the Danish-speaking 'børn' has been searched. There are manual cross-lingual mappings available for English and Swedish, and semi-automated mappings available for Danish and Mexican Spanish as well. By selecting Danish, a comparison chart of Danish bog and Swedishlovak and a Swedish Danish occurrence chart appears.

#### Compare with caution

Cross-lingual comparison is acknowledged taking advantage of the possibility of cross-linguistic comparisons. Because exact synonym is seldom found across languages, users must consider their research questions and hypotheses in determining what is to count as a mapping. Cognate status, phonological similarity, degree of semantic relatedness and other factors may be very important for specific properties. For this reason, CLEX provides a functionality by which the initial development/bidirectional data or a later research can carry his or her own dataset database on their own or other use. By default, American English serves as the ‘interlingua’ and comparisons between other languages are made via their links to it, but direct mappings between non-English languages is also possible.

### Add data to CLEX

Join CLEX

We invite other investigators who have developed an adaptation of the CDI to join in this project. All data to be transferred anonymously, only age and gender and possibly subcategories and the set of initial mappings of individual items to American English. We also anticipate adding existing datasets from specific atypical populations, such as Specific Language Impairment and Down Syndrome, when they are of adequate size and representativeness.

A major impetus for the development of CLEX has been the success of the Child Language Data Exchange System (CHILDES; MacWhinney, 2000) in promoting child language research. One crucial factor in the success of CHILDES has been the generosity of researchers in contributing data, and we hope that the CDI datasets can be used as a CLEX, like CHILDES, are we developing, and on platforms that permit exploration and analysis of the data. Another contrasting factor has been the inappropriateness and unavailability of CHILDES development and many others in the scientific discipline. We hope that researchers will contribute to CLEX not only CDI datasets, but also suggestions for useful tools which our facilitate research.

Technical information

Database: MySQL

Code language: HTML, CSS, JavaScript, SQL

### References


Rune N. Jørgensen, Center for Child Language, University of Southern Denmark

Philip Dale, University of New Mexico

Dorthe Blesnes, Center for Child Language, University of Southern Denmark

Larry Fenson, San Diego State University