**CDP++ Italian**

***Perry, C., Ziegler, J. C., & Zorzi, M. (in press). CDP++.Italian: Modelling sublexical and supralexical inconsistency in a shallow orthography. PLoS one.***

**Changes from other models**

CDP++.Italian works like all of the other CDP models.

There are some important differences:

1) First, to make life simpler for people with English keyboards, words that would have a diacritic on them have been changed to capitals. So café gets the new spelling cafE. If you type in letters that don’t exist in the model (i.e., the real letter), it will crash.

2) Running the full model (i.e., all words in the lexicon), which is done by unclicking fast on the menu for running words will take a very long time for each word. The model will more or less hang until it has finished computing. If you run batch files with this off with more than one word, then it will be faster than running the word separately because it only needs to load the lexicon once.

3) You will get some more information as output from the files that are saved -- Most notably to do with stress. For example, here is what the output will look like for the nonword “bimpiro”:

221 bimpiro+++++++++++++++++ bimpiro 0.52 0.46 0.00 1 0.650 0.544 0.000 0 3 1

The stuff you see, in order, is:

1) The number of cycles the model took to terminate

2) The word (plus some spaces – note that most of them are irrelevant since the program will only process up to 8 letters)

3) The phonology produced

4-6) The activation of the TLA network on the stress syllable nodes.

7) Which syllable was the stress winner in (1, 2, 3) where 1 = first syllable, 2 = second syllable, 3 = third syllable. You will have to work out penultimate or antepenultimate from the number of syllables in the word. For example, with a trisyllabic word, if the stress winner is ‘1’ it means the antepenultimate has won. If it is disyllabic, it means the penultimate has won.

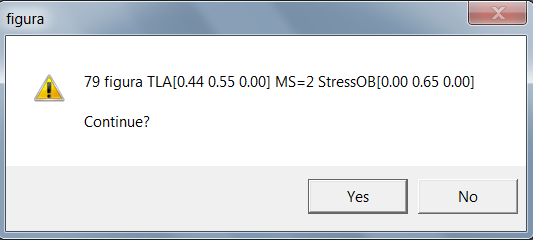
8-10) The activation of the stress output buffer in each syllable position

9) Whether the word/nonword is in the lexicon of the model

10) The number of orthographic vowels the model thinks there are (this will give you an idea of the number of syllables the model uses easily).

11) Whether the model finished within the number of cycles you gave it (I usually use a maximum of 250)

If you run a single word, you will also get more information



This gives you more or less the same information as previous (cycles,phonology,TLA levels, Stress Output Buffer levels)

As the model runs, you will get a file “WordInfo.txt” that will appear. This has the parsing which the parser used.

From here on in, everything is more or less the same as other versions of CDP. Therefore, I’ve just modified the manual

**NOTE THIS WAS INITIALLY THE CDP+ Manual. But the program works in the same way.**

**USER MANUAL**

**(rel. 11/06)**

**1. What is CDP+?**

CDP+ is a fully implemented computational model of reading aloud based on a dual route architecture built upon its predecessor, the Connectionist Dual Process (CDP) model developed by Zorzi and collaborators. It is the result of an incremental and nested modelling strategy, according to which a new model should build upon the strengths of its predecessors while eliminating their weaknesses. CDP+ therefore combines the best features of CDP (Zorzi et al., 1998, JEP:HPP) with those of the Dual Route Cascaded Model (DRC, Coltheart et al., 2001, Psych. Rev.) and the connectionist dual route model of spelling of Houghton & Zorzi (2003, Cog. Neuropsych.).

CDP+ is freely available as a research tool. No other use is permitted.

References for CDP:

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1131-1161.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273-315.

Perry, C., Ziegler, J. C., Braun, M., & Zorzi, M. (2010). Rules versus statistics in reading aloud: New evidence on an old debate. *European Journal of Cognitive Psychology, 22, 5*, 798-812. (German version)

Perry, C., Ziegler, J. C., & Zorzi (2010). Beyond single syllables: Large-scale modelling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology*, 61, 2, 106-151.

Zorzi, M. (2010). The Connectionist Dual Process (CDP) approach to modeling reading aloud. *European Journal of Cognitive Psychology,* 61(2), 106-151.

Perry, C., Ziegler, J. C., & Zorzi, M. (2013). A computational and empirical investigation of graphemes in reading. *Cognitive Science*, 37, 5, 800-828. (Updated the model so it parses graphemes)

Perry, C., Ziegler, J. C., & Zorzi, M. (accepted). CDP++.Italian: Modelling sublexical and supralexical inconsistency in a shallow orthography. PLoS one

Perry, C., Ziegler, J. C., & Zorzi, M (2014). When silent letters say more than a thousand words: testing computational principles across languages. *Journal of Memory and Language*, 72, 98-115.

Ziegler, J. C., Perry, C., & Zorzi, M (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369,* 20120397.

**2. Installing CDP+**

CDP+ is distributed as a stand-alone MS-Windows program that can be downloaded from the CDP+ web page located at the following url:

<http://ccnl.psy.unipd.it/CDP.html> or <https://sites.google.com/site/conradperryshome/>

The distribution file (CDPItalianFiles.ZIP) is a compressed archive that contains the executable (CDPItalian.EXE), seven configuration files, and this user manual. No installation is required. All files must be in the same folders as the zip will create (i.e., the config files must be in the ItalianRes directory).

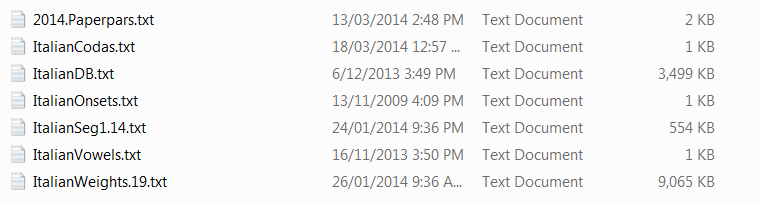
For any question or comment about CDP+, please contact by email one of the authors: Conrad Perry (conradperry@gmail.com), Jo Ziegler (ziegler@up.univ-mrs.fr), Marco Zorzi ([marco.zorzi@unipd.it](mailto:marco.zorzi@unipd.it)).

**3. Using CDP++.Italian**

**3.1 Loading CDP++.French**

To run CDP++.Italian, click on the CDPItalian icon.

If nothing pops up, it means something has gone wrong. This will happen if some of the files missing. These are the ones that should be in the ItalianRes directory.



**3.2 Running a Word**

If you want to run a word

1) go to RUN->Run Naming. A screen will pop up

2) Type in the maximum number of cycles to run the word. DO NOT USE VALUES OVER 300. Type in the word/nonword that you want to run. MAKE SURE THE WORD IS IN LOWER CASE apart from the letters changed for diacritics or the program will close e.g.,cafE for café.

3) Click OK. Wait... The phonology of the word will appear once it is named. The number of cycles needed to name the item is also displayed. In case it is not named within the given maximum number of cycles, a window with "Not named after XX cycles" will appear.

*Important notice: running CDP++.Italian is computationally intensive if you use the full lexicon (i.e., unlick the fast button – with it clicked it will just use those words in the lexicon that are an exact orthographic match with the word and its homophones). Processing a single word may take more than a minute depending on the computer speed.*

**3.3 Saving information**

The CDP++.Italian program allows to save the phonology/naming time of each item typed in

1) go to Saving -> QuickSave.

2) type in a file name (we suggest using the .txt extension)

Whenever the model produces a response, the output will now be written to this file (latency and pronunciation). Note that if the model does not name the item within the given maximum number of cycles, a "0" will appear after the item (otherwise, a "1" will appear).

Once you turn on the QuickSave option, you cannot turn it off except by reloading the program.

**3.4 Saving information II**

The CDP++.Italian program also allows you to save everything that happens in the network above the activation level of 0.02.

1) go to Saving -> SaveInterest.

2) type in a file name. Various markers will be appended to this.

This will save all activation of anything that is over 0.02. TotalSum and PhonSum are the summed lexical frequencies for the orthographic and phonological lexicons.

Use this feature only when necessary, since it will slow down processing a lot. It may also create massive files in some cases.

Once you turn on the SaveInterest option, you cannot turn it off except by reloading the program.

For stress, the characters #rL are used to represent what is going on in the 1st, 2nd, and 3rd syllable of the stress output buffer.

**3.5 Running a Batch File**

The batch feature allows you to run a list of items instead of typing them in individually.

To do this:

a) Make sure you specify a file name to save the results (go to Saving -> QuickSave) (you can also use “InterestSave”)

b) Go to Run -> Run Batch Naming

c) Choose the batchfile

d) Type any legitimate string into the word box, and a number into the number of cycles box (this currently does nothing, but you need to do it).

e) Press OK.

f) Wait…. Note that if you press “Cancel” the program will crash.

**3.6 Setting up a batch file**

The batch file needs

a) The word “naming” at the top

b) Each word/nonword followed by the maximum number of cycles

e.g.

naming

bildoro 250

bitora 250

bivata 250

-------------------------------------------------

CDP+ USER MANUAL

Copyright 2006 – C. Perry, J.C. Ziegler, M. Zorzi

-------------------------------------------------