LEARNING TO READ

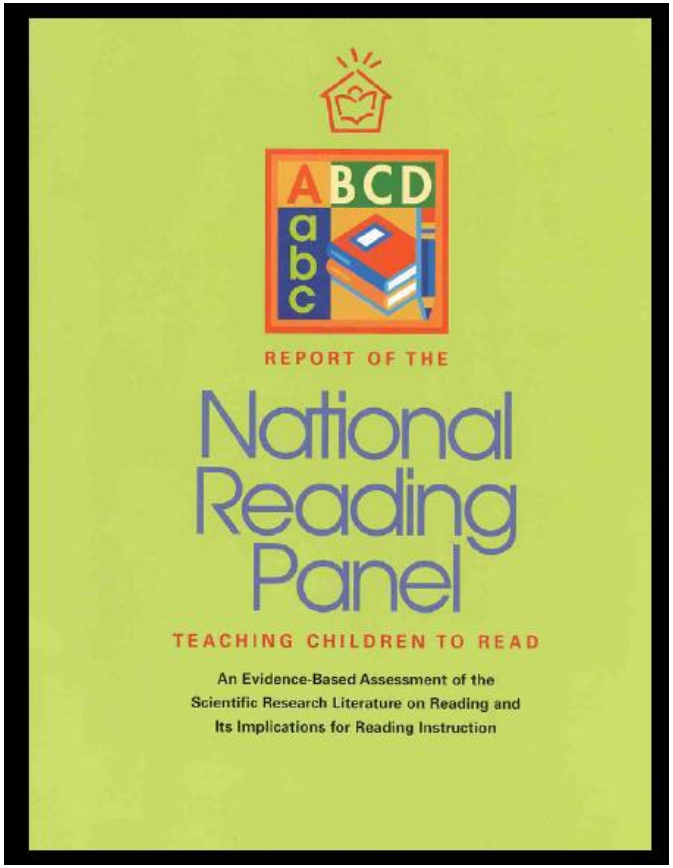
Marco Zorzi

Dipartimento di Psicologia Generale
Università di Padova

READING ACQUISITION

Phonological decoding (i.e., mapping letters onto sounds) is the *sine qua non* for reading acquisition

(Share, 1995, *Cognition*)

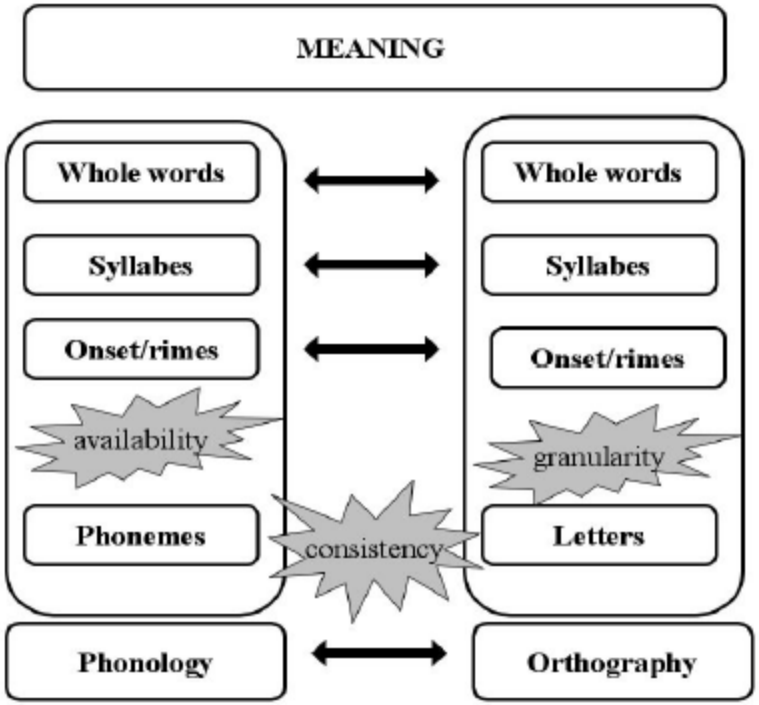


Three challenges for learners

I. **Availability:** not all phonological units are consciously / explicitly accessible before learning to read

II. **Consistency:** some orthographic units have multiple phonemic realizations (and vice versa)

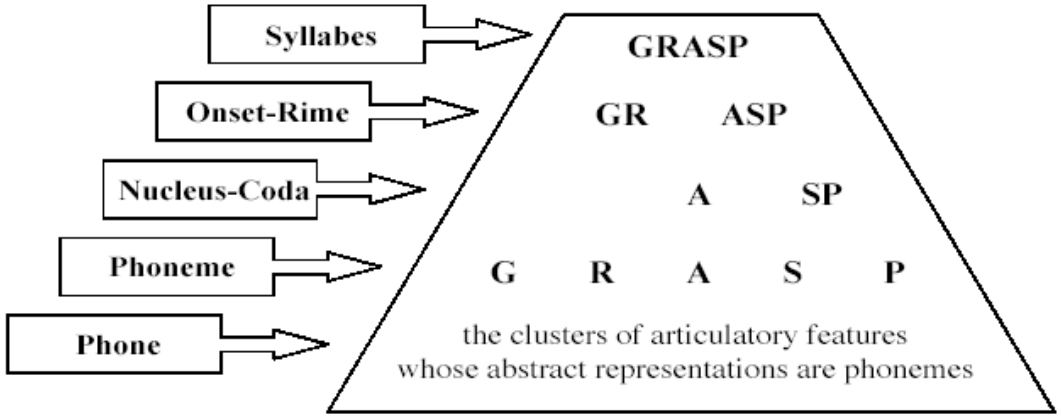
III. **Granularity:** there are more orthographic units to be learnt when access to the phonological system is based on larger «grain size» units (rimes, syllables, morpheme, etc.)



Ziegler & Goswami, *Psych. Bull*., 2006

Phonological development

- Measured through phonological awareness (PA)
- PA is strongly predicitive of reading acquisition (Bryant, etc.)
- Development of PA is sequential along a continuum from «shallow» awareness for large units to «deep» awareness for small units (Anthony et al., *JECP*, 2002; *RRQ*, 2003)



Data (% correct) from Syllable and Phoneme Counting Tasks in Kindergarten and First Grade Across Different Languages

	Kindergarten		First Grade	
Language	Syllable	Phoneme	Syllable	Phoneme
Turkish[1]	94	67	98	94
Italian[2]	80	27	100	90
Greek[3]	85	0	100	100
French[4]	69	2	77	61
English[5]	48	17	90	70

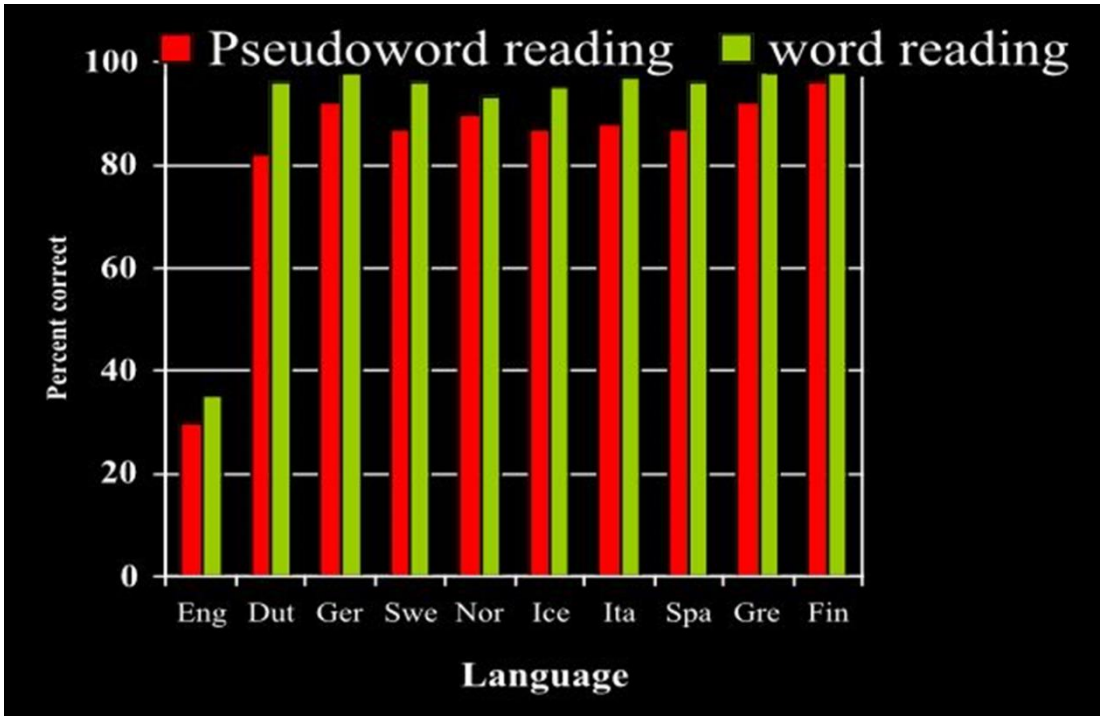
Note. 1 = Durgunoglu & Oney, 1999; 2 = Cossu et al., 1988; 3 = Harris & Giannouli, 1999; 4 = Demont & Gombert, 1996; 5 = Liberman et al., 1974

Cross-linguistic differences in learning to read

• Despite the similarity in the trajectory of phonological development, reading acquisition in different languages considerably variable.

• Three factors seem crucial for cross-language differences :

- *consistency* of the spelling-sound mapping
- *granularity* of orthographic representations
- *teaching* method



(Data from end of Grade 1; Seymour, Aro, & Erskine, 2003, *Br.J.Psych.*)

Consistency and granularity

- Units with smaller grain size tend to be more inconsistent that units with larger grain size (especially in less transparent orthographies like English)
- This induces children to develop a phonological decoding strategy that involves units of different grain sizes (for English, the rime; Goswami & Bryant, 1990)
- Presenting a block of non-words that are homogeneous in terms of granularity can facilitate reading with respect to a condition in which they are mixed. This holds only for a language in which both levels are processed (in English, letters vs. rime). German children do not show this effect (Goswami, Ziegler, et al., *App.Psycholing.*, 2003)
- Also the involvement of lexical (whole-word) strategies depends on consistency. The advantage in reading pseudo-homophones is present in English but not in German children (Goswami, Ziegler, et al., 2001, *JML*)

The role of teaching stategies

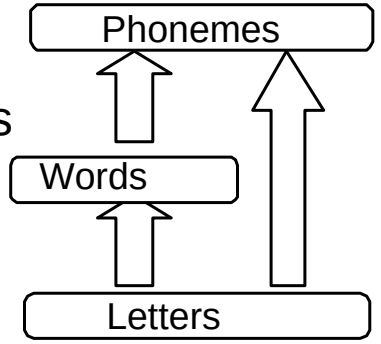
• In non-alphabetic orthographies (e.g., Chinese) children must memorize a great number of characters, a process that takes up to 5 years. In China an alphabetic writing system (*pinyin*) is taught first, which is then used on the side of the Chinese characters to be learned.

• In alphabetic languages, teachers usually start to teach reading starting from single letters. The child learning the mapping between letters and sounds acquires knowledge about phonemes. Experience with written language changes the nature of phonological representations.

• Early teaching of reading, focused on a phonetic method, has been proposed as a standard in the UK (DfEE, 1998). Teaching based on phonemes is particuarly effective in languages with consistent spelling-sound correspondences. It is unclear whether explicit teaching of strategies based on larger units really helps learning in languages with inconsistent orthographies (USA National Reading Panel, 2000; Walton, *SSR*, 2002).

The Connectionist Dual Process (CDP) approach to modeling reading aloud

Learning to read aloud demands decomposition of the task into ***phonological assembly*** *vs* ***phonological retrieval***



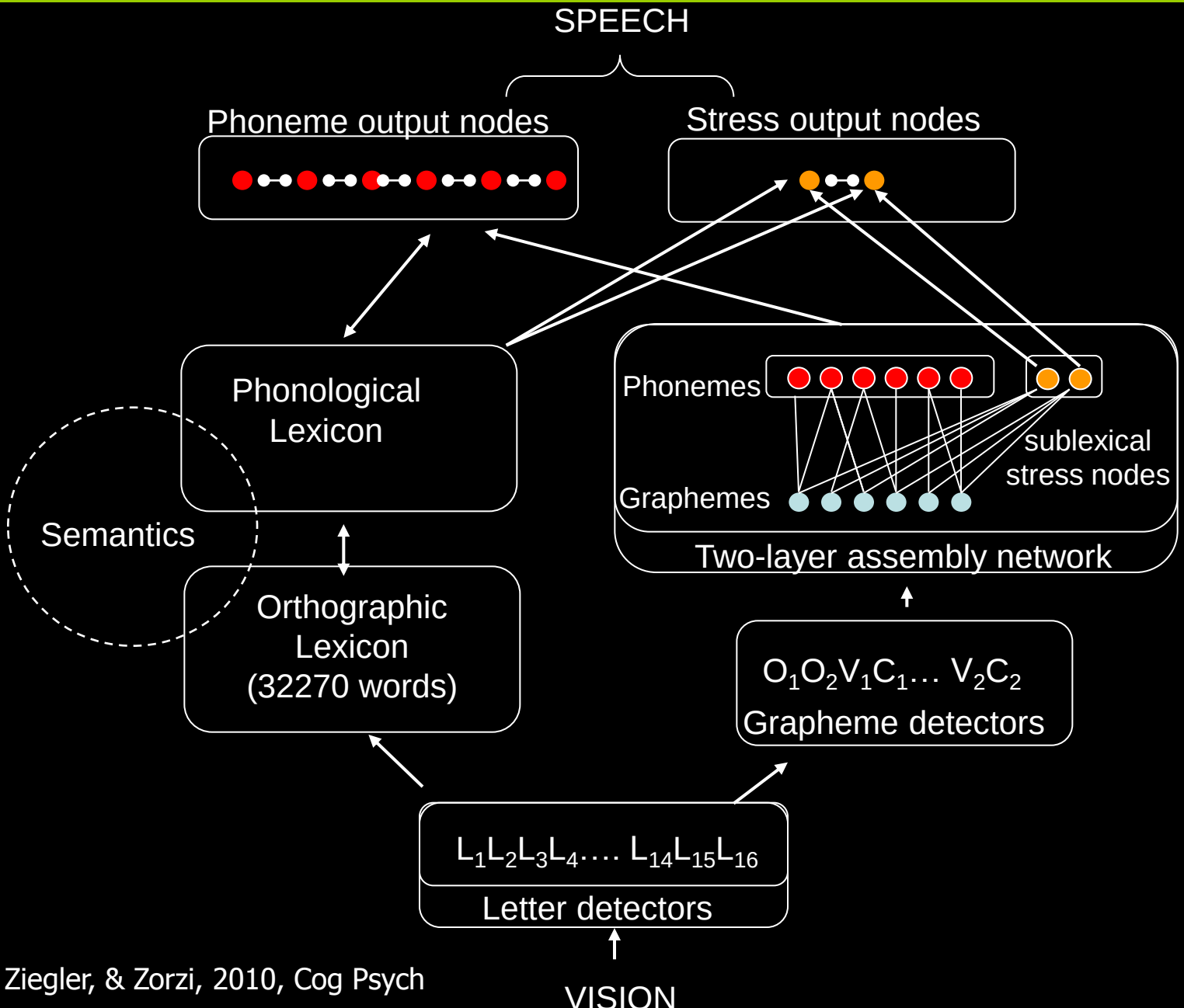
• **Phonological assembly (decoding)**: A pathway where letters (graphemes) become directly associated with phonemes

• **Phonological retrieval**: A pathway where the computation of phonology is mediated by whole-word (lexical) representations

CDP is a dual route theory but it differs from the "classic" dual-route model where phonological decoding is attributed to a slow, controlled process based on explicit symbolic rules (e.g., Coltheart's DRC).

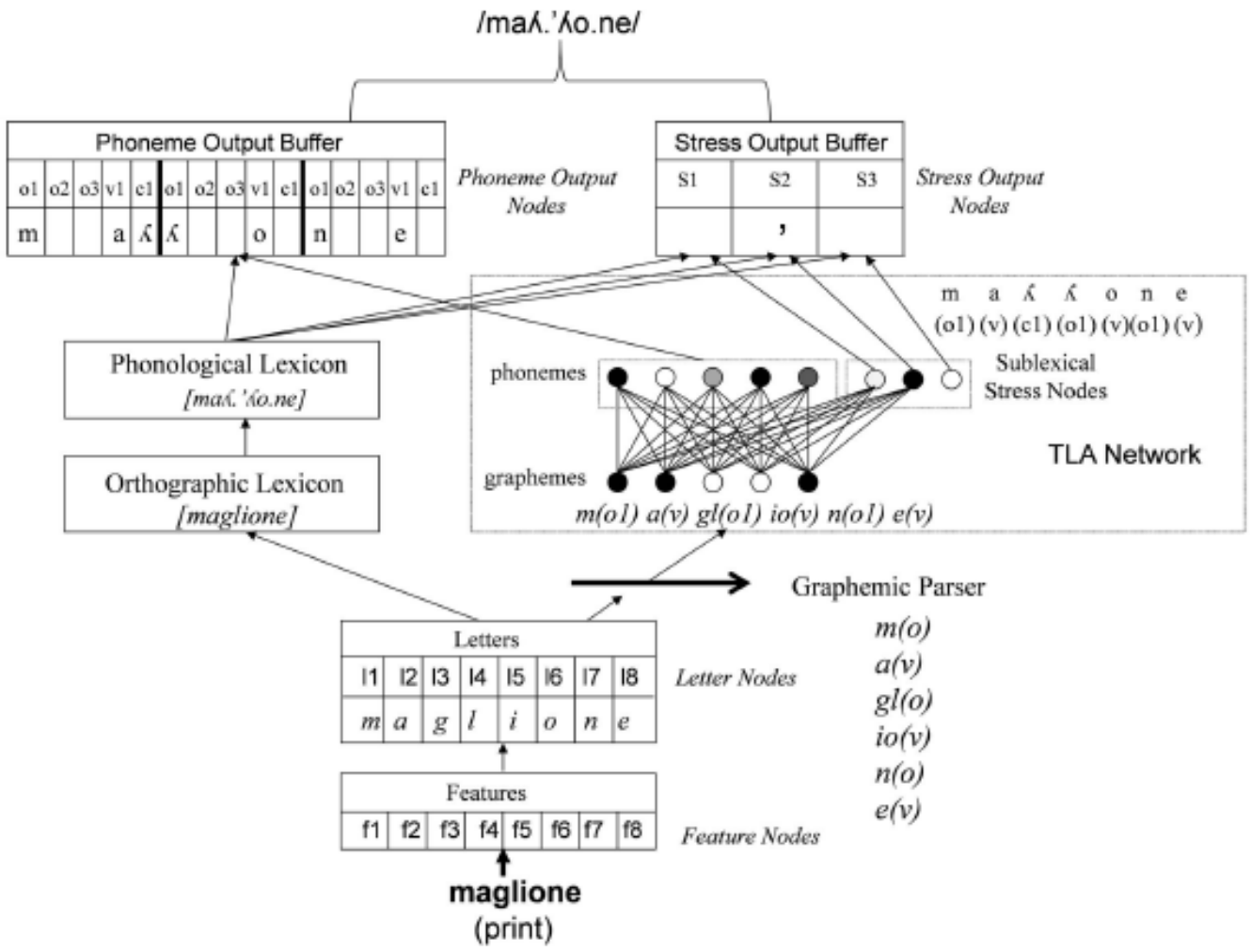
A key component in CDP is the **decoding network** (or *two-layer assembly* network), a simple linear system that learns the statistically most reliable spelling-sound mappings.

The CDP ++ model of reading aloud



Perry, Ziegler, & Zorzi, 2010, Cog Psych

CDP++.Italian



Perry, Ziegler, & Zorzi, 2014, *Plos One*

The “learning rate effect”

Learning to read a relatively irregular orthography like English is harder and slower than learning a relatively regular orthography like German.

Can this be explained by connectionist models of reading?

(Hutzler, Ziegler. Perry, Wimmer, & Zorzi, 2004, *Cognition)*

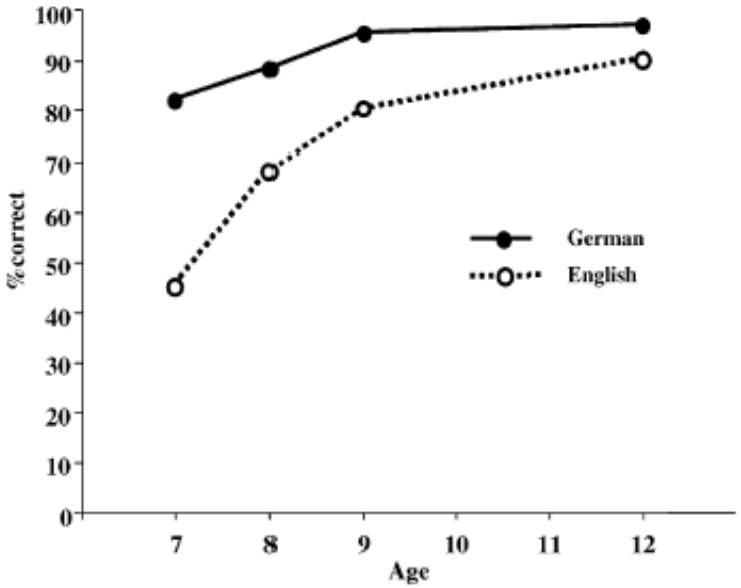


Fig. 1. Prototypical data pattern illustrating the learning rate effect that can be observed in literally every cross-language comparison involving an orthographically consistent (e.g. German) and the relatively less consistent English orthography. Data reproduced from Frith et al. (1998).

The comparison between English and German can be based on of the same set of (non)words

(Ziegler et al., 2001, *Psy. Sci*.)

• Learning corpora for English and German are matched for number of words and average frequency

• 80 test nonwords are identical in the two languages (fot–Fot, lank–Lank, plock–Plock, etc.). Length (4-6 letters – 20 for each length), consistency and orthographic neighborhood are matched

• The model shows a German advantage over English but not correct shape and timing

• *The model's sensitivity to the statistical regularities of the spelling-sound mappings is not sufficient to capture the exact pattern of reading acquisition*

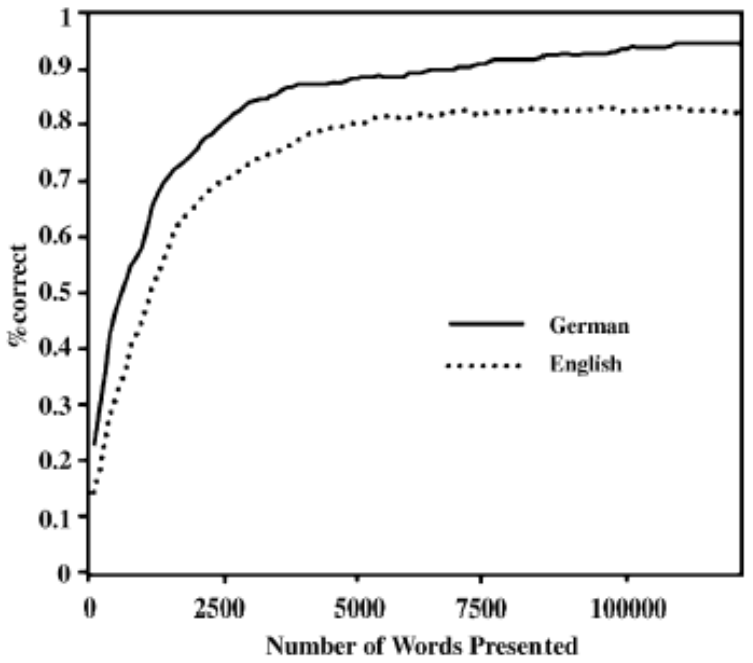


Fig. 3. Nonword reading performance of the German and English implementation of Zorzi et al.'s (1998a) two-layer associative network during the course of training.

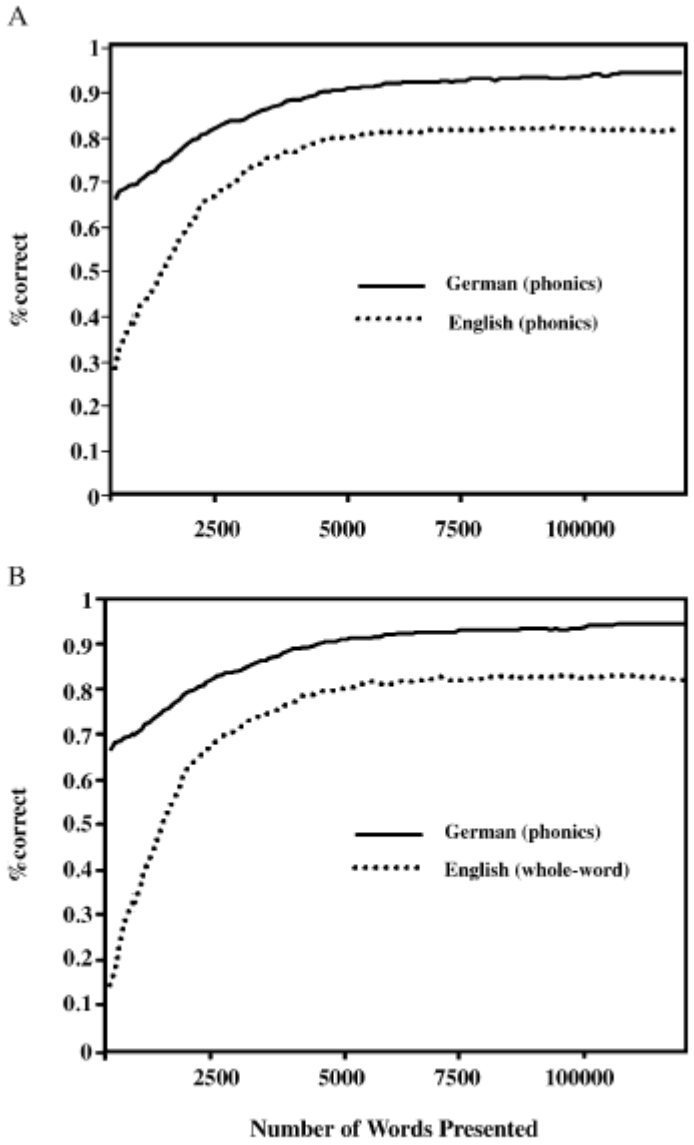


Fig. 4. Nonword reading performance of Zorzi et al.'s (1998a) two-layer associative network when both implementations are pre-trained using a phonics regime (A) and when the German phonics approach is compared to the English whole-word approach (B).

The role of teaching method: phonics vs. whole-word

Phonics: pre-training on spelling-sound correspondences (E: 66; G: 64) derived from phonics teaching programs (e.g., *Jolly Phonics,* Lloyd, 1999)

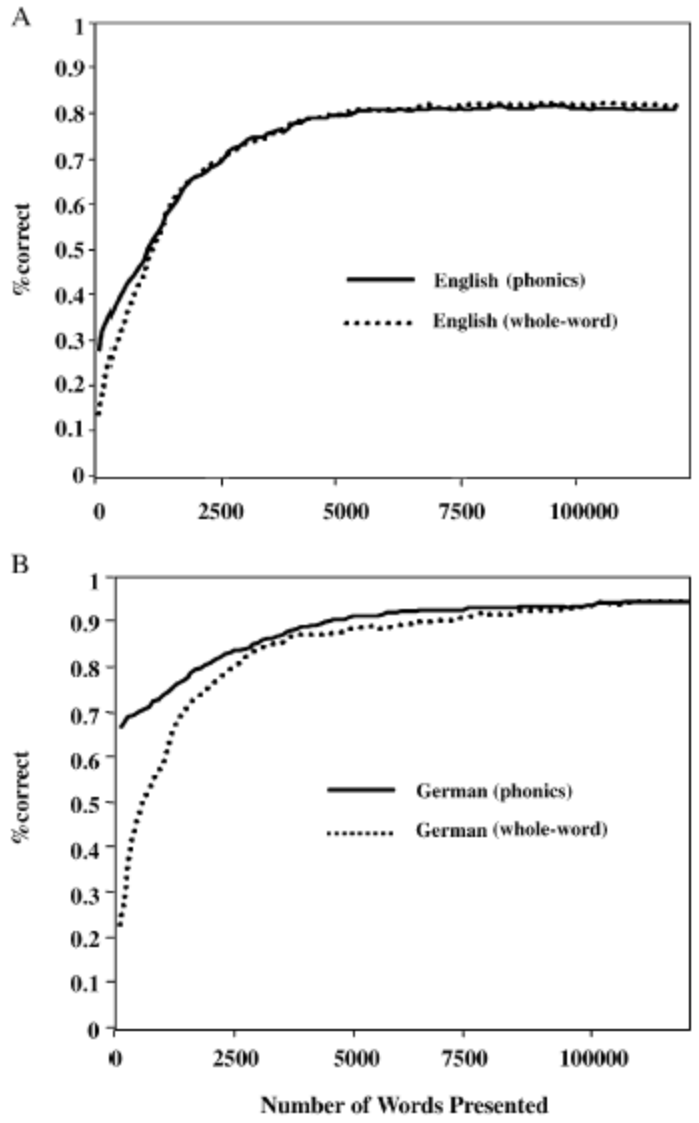


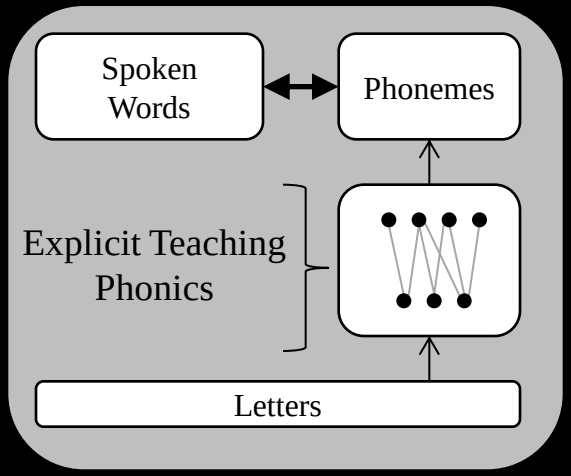
Fig. 5. Benefits due to phonics pre-training of the English and German versions of the two-layer associative network (A,B, respectively).

Effect of teaching method

• The benefit of phonics teaching is modulated by the consistency of the mapping (i.e., more marked for German; cf. Landerl, 2000, JPE).

The self-teaching mechanism

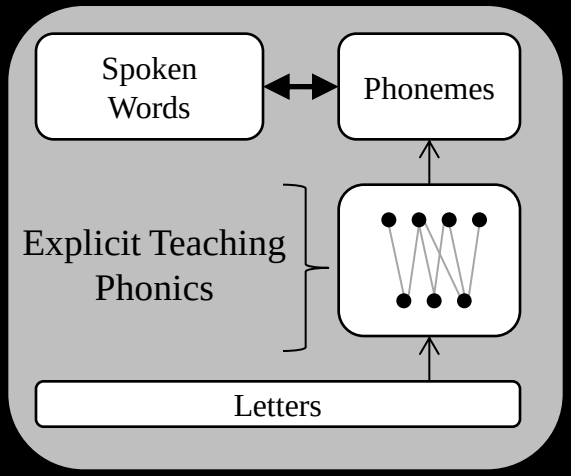
Explicit instruction



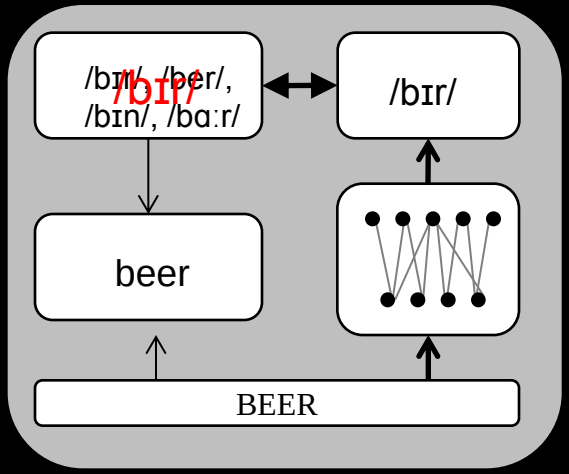
- Phonological lexicon is in place before learning to read
- Children learn about letters
- …and how they map onto phonemes

Ziegler, Perry & Zorzi, 2014, *Phil. Trans. Royal Soc.*

The self-teaching mechanism

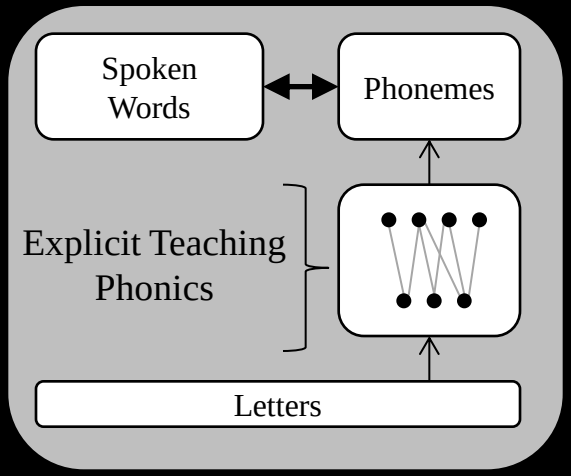


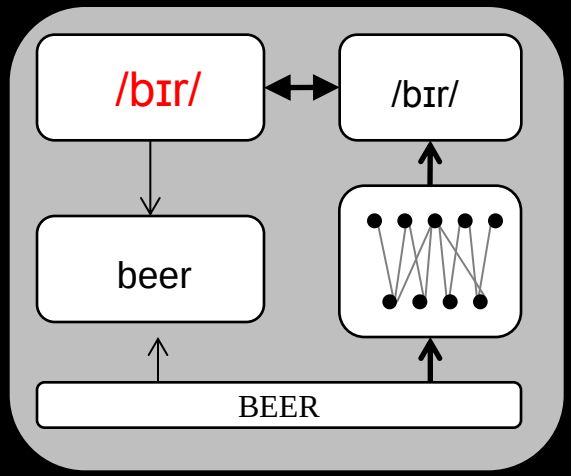
Decoding allows the child to retrieve the phonological form of the word and build an orthographic representation



Phonological decoding and orthographic learning

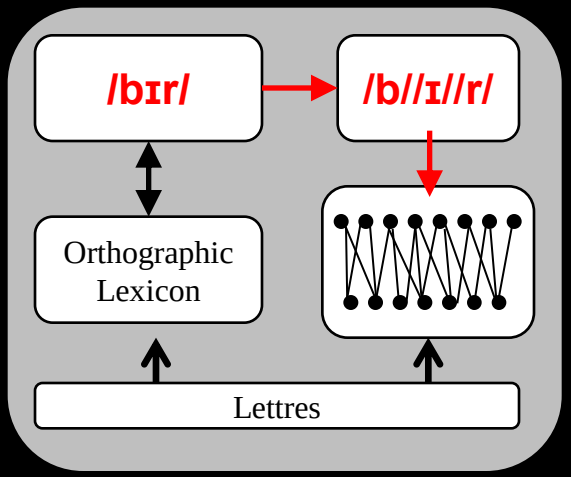
The self-teaching mechanism



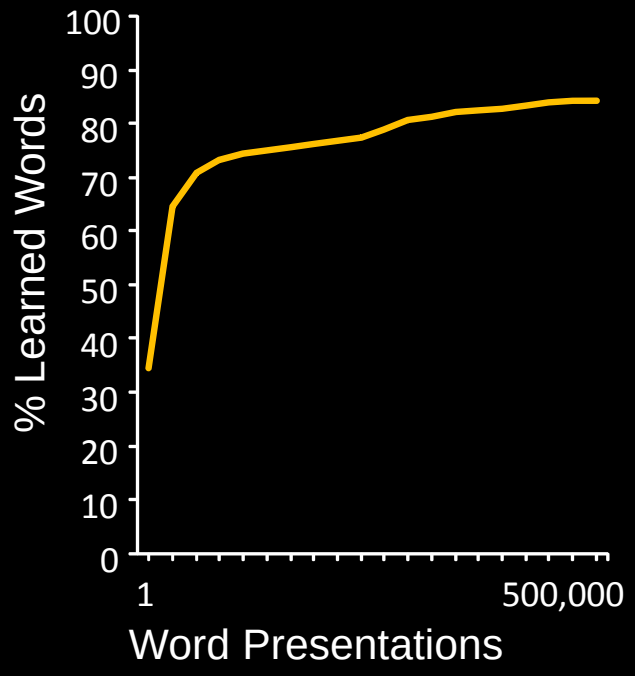


Consolidation and automatization

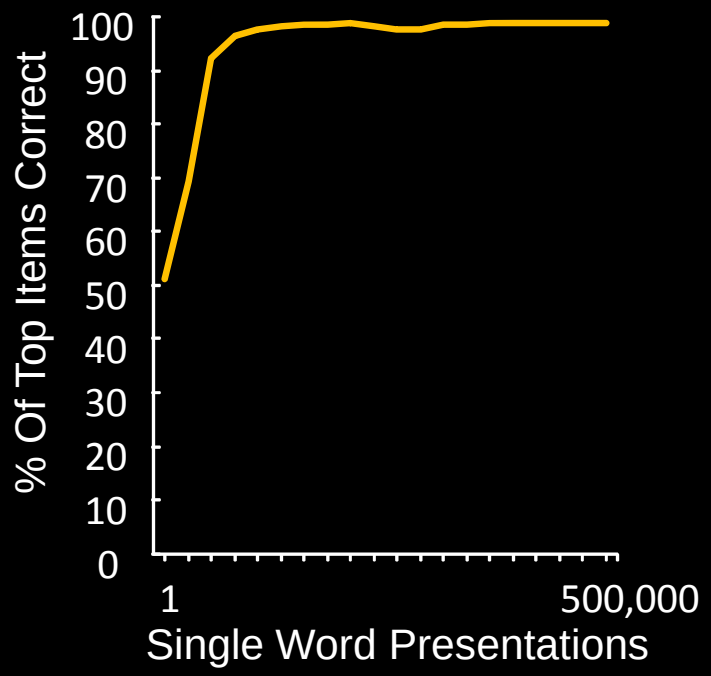
Self-teaching: the decoded word is used as internally generated teaching signal to improve decoding



Learning to read English

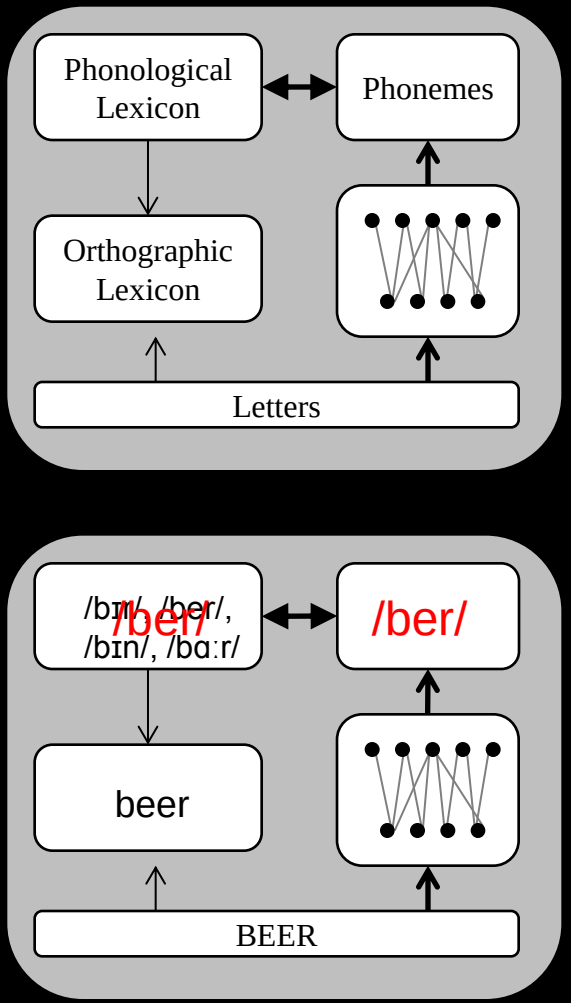


80% of English words can be learnt through decoding and self-teaching

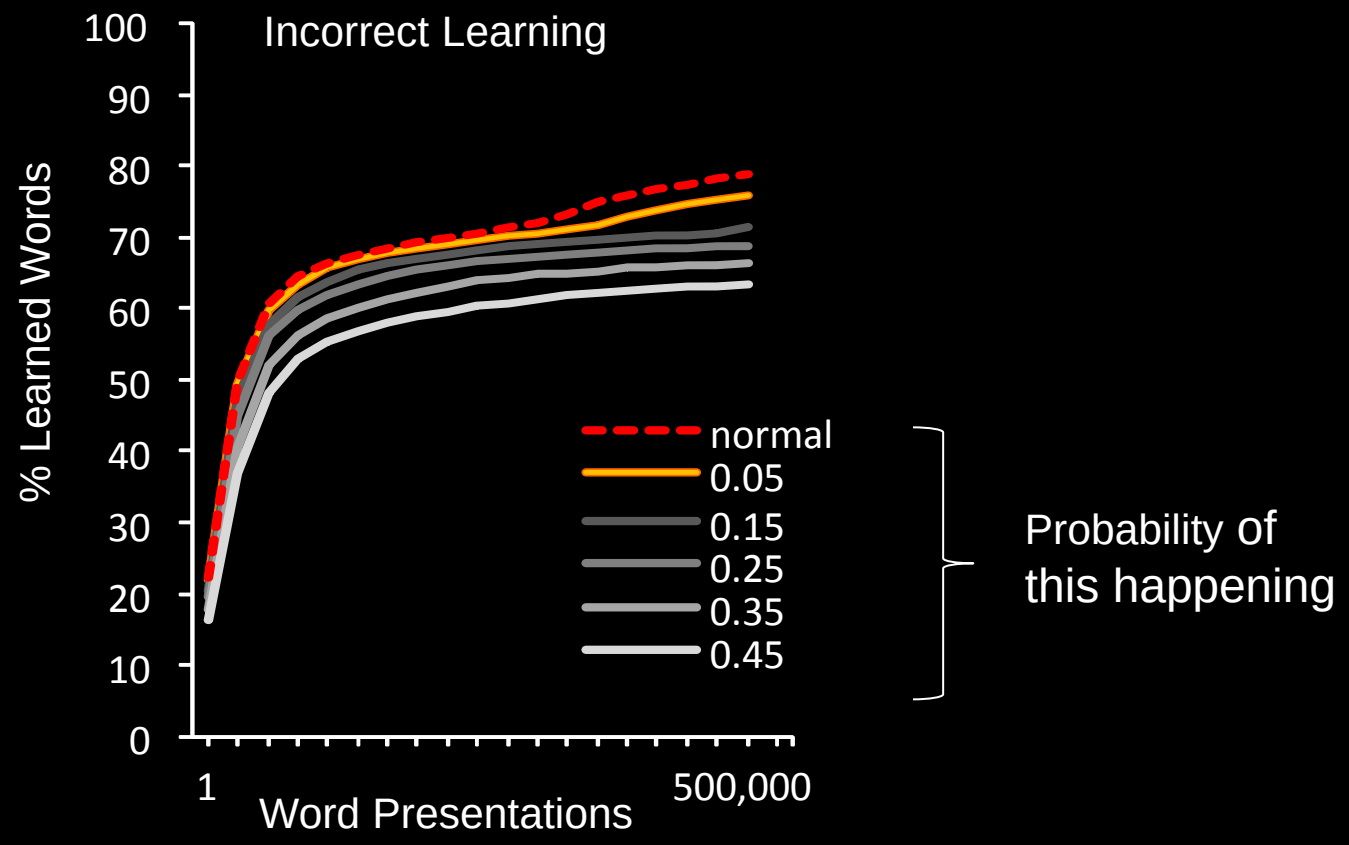


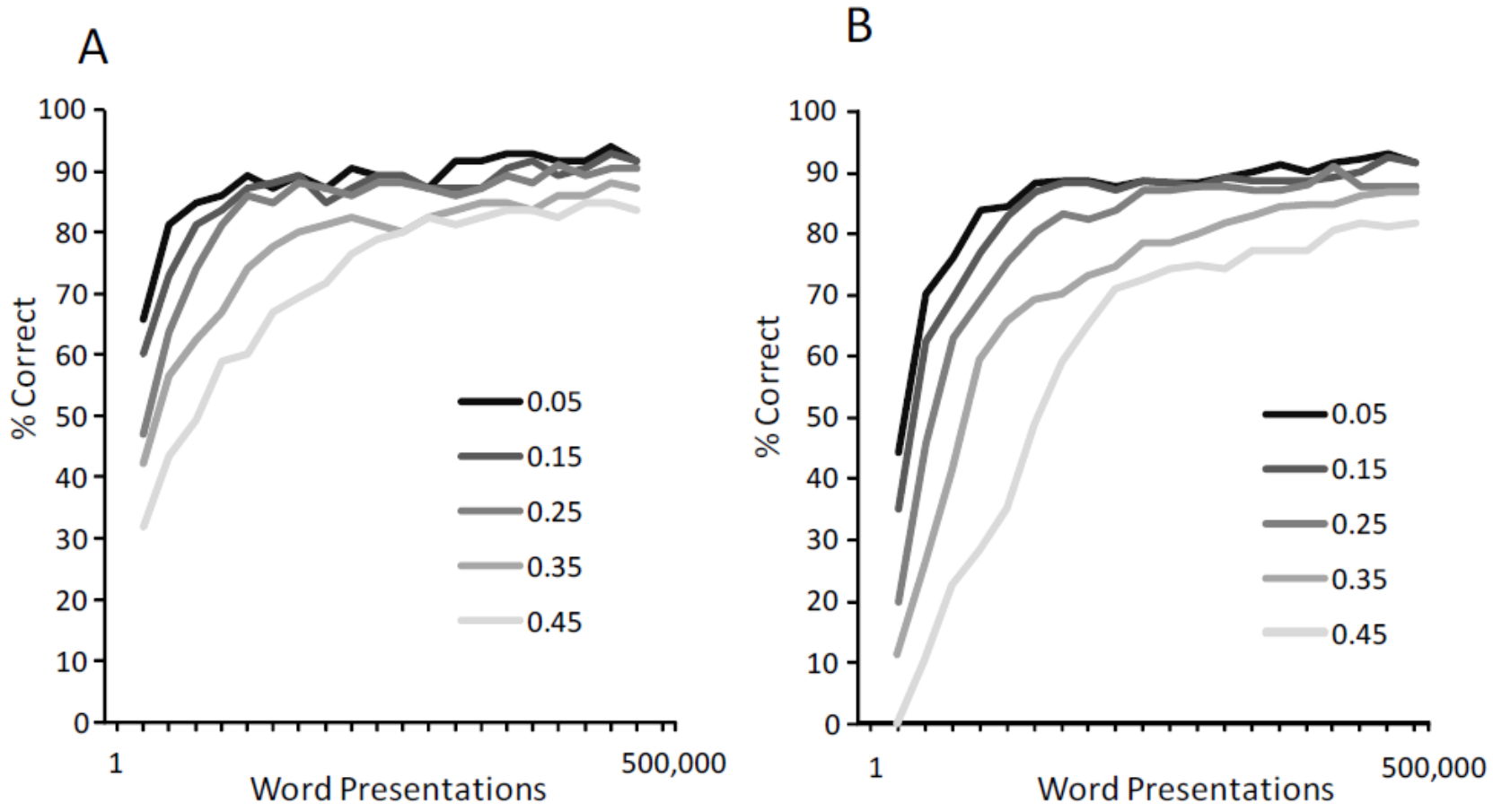
Very quickly the correct word is the most active one in the phonological lexicon

What happens when an incorrect word is activated?



Incorrect learning trials do not hamper overall learning





Can the model read novel words? Generalization performance of the model on two set of nonwords (easy nonwords, left; difficult nonwords, right) The nonword sets were repeatedly tested during the course of learning to read.

Interim summary

• Connectionist learning models must be sensitive not only to statistical regularities of the spelling-sound mapping but also to the teaching method to adequately capture the pattern of reading acquisition

• Phonics teaching is particularly effective in languages with consistent orthographies. It is controversial whether explicit teaching of strategies based on larger units (e.g., rimes, syllables) would help learning to read inconsistent orthographies (National Reading Panel, 2000; Walton, *SSR*, 2002). This issue could be addressed in simulation studies. However, CDP+ simulations show that parsing/alignment of graphemes into a syllabic template is the key to efficient learning of phonological decoding skills

• A teaching method focused on phonemes is the standard in the UK (DfEE, 1998), USA (National Reading Panel, 2000) and France (from 2006/07). The use of global/holistic methods has no scientific support.

• Phonological decoding bootstraps reading acquisition and the self-teaching mechanism is crucial for orthographic / lexical learning