Do computational models of reading need a bit of semantics?

Remo Job
Università degli Studi di Trento, Italy
And
Claudio Mulatti
Università degli Studi di Padova, Italy

Abstract. Coltheart, Rastle, Perry, Langdon, & Ziegler (2001) claim that "the psychology of reading has been revolutionized by the development of computational models of visual word recognition and reading aloud". They attribute this to the fact that a computational model is a computer program – an algorithm – "that is capable of performing the cognitive task of interest and does so by using exactly the same information-processing procedures as are specified in a theory of how people carry out this cognitive activity" (Coltheart et al., 2001; page 204). According to this view, the computational model is the theory, not a simple instantiation of a theory.

In this paper we argue that computational models of reading have indeed helped in dealing with such a complex system, in interpreting the phenomena underlying it, and in making sense of the experimental data. However, we also argue that it is crucial for a model of reading to implement a computational semantic system that is as yet a missing component of all computational models. We provide two reasons for such a move. First, this would allow explaining some phenomena arising from the interaction of semantics and lexical variables. In this section, we will review the following empirical findings: faster response times to polysemic words (e.g. Hino & Lupker, 1996) and slower response times to synonyms (Pecher, 2001); the leotard (Rodd, 2004) and turple effects (Forster and Hector, 2002); and the asymmetry of the neighbourhood density effect in free and conditional reading (Mulatti & Job, 2006). Second, such an "enriched" model would be able to account for a richer set of tasks than current computational models do. Specifically, it would simulate tasks that require access to semantic representation to be performed, such as semantic categorization and semantically-based conditional naming.

We will present a computational instantiation of a semantic module that accounts for all the described phenomena, and that has helped in generating predictions that guides on-going experimental activity.

1. What a computational model is (for us)

Without theories (models[1]), our ability to understand the great deal of data generated by experimental observation would be sporadic and limited to isolated facts or cases. Models of cognitive processes, then, constitute frameworks which help scientists in dealing with such complex systems, in interpreting the phenomena and in making sense of the experimental data. A model of a cognitive process describes and explains that cognitive process. For example, a model in the visual word recognition field – under the assumption that it is an appropriate model – describes and explains the processes underling reading.

Ontogenetically, models are first expressed verbally[2]. In the visual word recognition domain, a verbal model describes and explains processes trough the utilization of natural language (sentences), graphical supports (flowcharts – the so called boxes and arrows models), or both. A verbal model, then, is a qualitative one. Shortcomings of purely verbal theories are vagueness, ambiguity and imprecision (Broadbent, 1987)[3], reluctance to falsification, and confusability, in the sense that a qualitative description is not easily distinguishable from other qualitative descriptions (Massaro, 1992). Last but not least, verbal model are too easily adaptable, extendible, to new data, even if inconsistent (provided the inconsistencies are "comfortable"; Jacobs & Grainger, 1994). However, to regard verbal models as always inadequate would be an error, since they do present with positive aspects: "[they] attract the expression of creative ideas, when the database is still too sparse to reasonably constrain more formal models.

---

[1] The terms *theory* and *model* are treated as interchangeable.

[2] Generally, this statement is false. As Jacobs and Grainger (1994) pointed out in the field of word recognition, "two distinct approaches to model construction emerge from the literature […]. The first, which may be coined the *gardener's approach* (or 'the model is not the theory'), can be caricaturised as consisting in 'growing' a model or network that mimics in some respect a human cognitive function, without necessarily having an explicit theory of that function […]. The second strategy could be coined the *architect's approach* (or 'the model is the theory'). In line with the central dogma of cognitive science (cf. Chomsky, 1965), some continue to argue that it is the right approach to start with a fully specified theory (based on general principles) and then (if one wishes) to implement it as an algorithmic model". The gardeners first develop a computational model that works and then they develop a theory compatible with the model. The architects first develop a theory compatible with the data and then they develop a computational model to test the theory. If the focus is restricted only to the architect's approach, the dead end is overcome since the statement gets true.

[3] Broadbent (1987) was able to demonstrate how a single algorithm could explain four patterns of results which were previously explained by four different verbal models, so giving a clear example of the explanatory inadequacy of verbal models.

[they also] attract the organization of results coming from a broad variety of tasks" (Jacobs & Grainger, 1994, page 1312).

When a model is no longer a sketch of a cognitive process but, rather, it describes and explains the procedures involved in that cognitive function in a greater detail, the model is often (that is, when possible) translate into a *computational*[4] *model*. A computational model simulates a mental function by implementing a theory. More precisely, a computational model is a computer program – an algorithm – "that is capable of performing the cognitive task of interest and does so by using exactly the same information-processing procedures as are specified in a theory of how people carry out this cognitive activity" (Coltheart et al., 2001; page 204). Therefore, the computational model *is* the theory, not a simple instantiation of a theory (but see Norris, 2005).

Coltheart et al. (2001) go on to say that "the psychology of reading has been revolutionized by the development of computational models of visual word recognition and reading aloud", which is agreeable since computational modelling has obvious benefits. Firstly, in order to be implemented in the form of a computational simulation a model needs to be represented in an explicit form which imply a level of specification that typically eludes verbal theories. Secondly, the modeller needs to solve issue at a local level he/she may not even be aware of. Thus, the attempt to develop a computational model may interact with the theory itself, giving raise to a reciprocal improvement: the translation itself can be a productive process since this operation can uncover gaps or inconsistencies. Furthermore, since computational models are built on mathematical laws, the principles of their operations are explicit; this rigorous declaration of a theory can enable more accurate communication of ideas and reduce the scope for misinterpretation.

Jacobs & Grainger (1994; page 1312) listed a few "possible drawbacks of algorithmic models [which] are the dangers that they fossilize thinking and restrict creativity more than verbal models; that they focus the model builder's attention too much on […] implementation details that are irrelevant and thus obscure the discovery of general principles; that, in absence of a computational theory, they are not more than mimicry (Marr, 1982); or that they cannot explain much if they still have to be explained themselves (Olson & Caramazza, 1991)".

The computational model itself can become the subject of investigation. A computational model simulates cognitive function. Such simulation not only enables theories to be put on the test, but it provides tools for

---

[4] Marr (1982) reserved the term *computational* for highest level description, whereas he called *algorithmic* the simulation models. Here, no distinction between the terms algorithmic and computational are made, so that the terms computational and algorithmic are interchangeable.

investigating the theory itself and making likens between alternative theories. A computational model can be used to explore a theory in ways that would otherwise be either beyond the scope and the possibility of experimental investigation – as, for example, the quantitative estimation of the relative weight of competitive procedures in determining the output of a function – or too complex to be faced purely from the behavioural data.

Moreover, computational models help in generating predictions that can guide future experimental activity. The practice of modelling, generating predictions and testing these predictions, leads to improvements in the model itself, and doing so improves knowledge about the cognitive function mimed by the model.

In the next sections, we will briefly address some issues arising in the study of reading aloud single words. We will then describe a computational model of reading aloud and visual word recognition. After that section, we will show how computational modelling should be used *in practice* by deriving some predictions from a model and testing them. Finally, we will discuss how the model accounts for the body of empirical data.

2 Reading aloud: Some issues

Word reading is a complex cognitive operation requiring, at least, a mechanism that maps print into semantics and (or) into phonology. Not surprisingly, theories of word recognition usually posit some sort of lexical path (print to sound) and some sort of semantic path (print to meaning), although that distinction might not be so explicit and many researchers (Lukatela & Turvey, 1994a, 1994b) believe that meaning can only be accessed once the phonological representation has been retrieved (print to sound to meaning). Even if different theories might not – and often do not – converge on what the minimal set of assumptions needed is, what the nature of the computations and of the representations is, or what the structures of the processes involved look like, there is an aspect that is common to all of them: whereas the process of deriving sound from print is described in details, the process of accessing the meaning is usually under-specified in terms of both the representations involved and the procedures operating on those representations (e.g. Plaut, McClelland, Seidenberg, & Patterson, 1996; Coltheart et al., 2001). The causes of this deficiency have to be tracked back to the nature of the semantic representations itself, which is fleeting and hard to capture, to entrap into a describable format, despite the considerable amount of empirical evidence and notable recent theoretical contributions (e.g. Cree & McRae, 2003; Sartori & Lombardi, 2004; Vigliocco, Vinson & Lewis, 2004).

This vagueness in the verbal descriptions of the semantic system directly reflects into the computational models that from those theories are derived. Although currently available computational models of reading and visual word recognition do a great job in explaining/simulating (various portions of) the set of data reported in the literature (e.g. written frequency, letter length, orthographic neighbourhood size, orthographic neighbourhood frequency, regularity, position of irregularity, body-rime consistency) they do so without implementing any semantic system. Noteworthy, all the effects that they explain/simulate are phenomena that can be ascribed to the operations of the lexical system: since they do not implement any semantic system, they cannot account for phenomena arising within the semantic system itself or due to the interchange of information between the semantic system and other systems, e.g. the orthographic input system.

To avoid the problems arising while trying to model the semantic system from a purely theoretical starting point, in the work here reported we choose a different approach. We select (on the base of both personal preferences and explanatory power) a suitable computational model of word recognition, and look at the implications ensuing from adding a minimal semantic system. This approach has some immediate benefits. It allows testing the plausibility of the assumption underlying the model in contexts different from those the model was originally developed for. It also allows testing the minimal computational apparatus needed to simulate (at least some) semantic effects. Moreover, it allows a better understanding of the dynamic of processing of the already implemented components, such as the orthographic and phonological systems.

We will now briefly describe the computational model we selected, the *Dual Route Cascaded* model (DRC; Coltheart et al., 2001).

3. The DRC model

Architecture of the model.

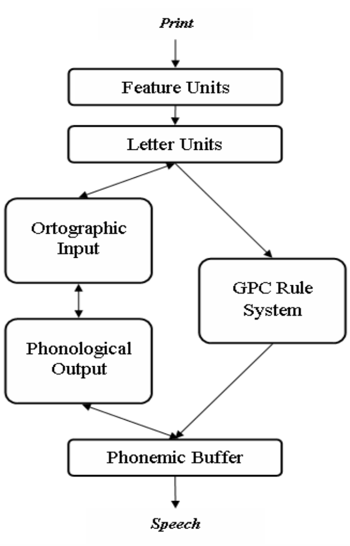
The general architecture of the DRC is outlined in Figure 1. It can be split into two parts: a. parallel search within a dictionary, the *lexical* routine; b. serial conversion of graphemes into phonemes, the *non-lexical* routine.

The Feature and Letter Identification levels, as well as the Phonemic Buffer, are shared by the two routines. Each unit in the Feature level represents one of a letter's features, each unit in the Letter level represents one letter of the alphabet, and each unit in the Phonemic Buffer represents one phoneme of the target language.

The *lexical* routine's specific components are (a) the Orthographic Input Lexicon and (b) the Phonological Output Lexicon. The lexicons consists of lexical entries which are localist nodes that represent each word known to the model in terms of its spelling (in the orthographic input lexicon) and sound (in the phonological output lexicon). The lexical routine works in parallel.

The model assumes: interactive activation between levels - with compatible units activating each other and incompatible units inhibiting each other, with the exception of the connections between the units in the orthographic and phonological lexicons, which are only excitatory; inhibition within levels – units belonging to the same level inhibit one another through inhibitory lateral connections.

Presenting a word to the model causes the activation of the visual feature units involved. Subsequent to the Feature level, activation is cascaded across all remaining levels. The features activate Letter level representations, which activate units in the Orthographic Input Lexicon which in turn activate units in the Phonological Output Lexicon. Activation then spreads to the Phonemic Buffer. The word is said named when activation of the rightmost phoneme of the word's phonological representation in the Phonemic Buffer reaches a pre-specified criterion.



Non-words can be named by virtue of the *non-lexical* routine, a mechanism that converts graphemes into phonemes through a procedure that apply grapheme to phoneme conversion (GPC) rules of correspondence. There are three kinds of rules: Single Letter rules, which apply when a single letter maps into a single phoneme; Multi-Letters rules, which apply when a group of letter maps into a single phoneme; Context-Sensitive rules, which apply when preceding or following letters consistently determine the pronunciation of a given grapheme. This mechanism operates serially, left to right, on the output from the Letter level and activates phonemes in the Phonemic Buffer. Letter information becomes progressively available to the non-lexical route. At the first cycle, no letters are available. After a constant number of cycles (10) the first letter is assembled into a phoneme. After this, every 17 cycles another letter becomes available to the routine, until all the letters have been processed or the criterion has been reached.

The amount of activation and inhibition sent between and within levels as well as the relative weigh of the two routes in assembling the stimulus' phonology is controlled by a set of 32 parameters. Moreover, using a single parameter set the DRC simulates 18 effects singled out in reading English.

Processing fashion and architecture of the model are described in greater detail in Coltheart et al. (2001).

Spread of activation.

Among the components of the lexical route, activation spreads in a cascaded fashion (*cascaded processing*, McClelland, 1979). In models that operate by thresholded processing, as for example the logogen (Morton, 1961), the processing going on in any module does not begin to affect subsequent modules at an early point in processing; activation in only passed on to the later modules after a threshold is reached in the earlier module.

In models that operate by cascaded processing, as the DRC, there are no thresholds between modules; as soon as there is even the slightest activation in an early module this flows on to later modules. This way of spreading activation is, within the DRC framework, crucial to simulate a few effects such as the effect of orthographic neighbourhood size in word reading aloud (Andrews, 1989, 1992; Sears, Hino, & Lupker, 1995).

4. An emergent phenomena: the orthographic neighbourhood size effect

The orthographic neighbourhood of a given word is the set of words that can be created by replacing one letter a time of that word. Different words can have neighbourhood of different sizes (e.g. CART: calf, calm, card, care, carp, cars, cast, cert, coat, curt, dart, hart, mart, part, tart, wart; FROG: flog, from, grog). Given a word, the size of its orthographic neighbourhood influences the time required to read it, indeed, as the number of neighbours increases, reading times decreases (Andrews, 1989, 1992; Sears et al., 1995). Within the DRC framework, this effect naturally emerges from its architecture and processing fashion, indeed "[…] cascaded processing in the model allows [words] to activate orthographically similar words in the orthographic lexicon, and this activation then feeds down to the phonological lexicon and finally to the phoneme system. Because generally the [neighbours'] units that became activated [in the lexicon] share phonemes with the stimulus, phonemic activation generated from the lexical route […] should facilitate stimulus [reading]" (Coltheart et al., 2001).

5. Two effects

Because of the cascaded processing, a word presented to the model activates all the orthographically similar words. Although the models does not implement any semantic module, cascaded processing allows us to make a rather straightforward prediction: since a word activates all the orthographically similar words in the orthographic lexicon, it activates their semantic representations as well. Two studies seem relevant here, one conducted by Rodd (2004), one by Sears, Hino and Lupker (1999).

Rodd (2004) presented her participants with words, one at a time. They had to perform a *semantic decision*, that is they had to decide whether the words were the name of an animal or the name of something else by pressing one of two buttons. Among the stimuli she used there were name of non-animal things (*leotard*) that had the name of an animal as neighbour (*leopard*). She showed that participants took longer to reject words with an animal name as an orthographic neighbour with respect to words without that sort of neighbours. It must be concluded that *leotard* activated the semantics of *leopard* enough to interfere with the semantic decision process.

Sears, Hino and Lupker (1999, experiment 3; see also: Forster & Shen, 1996; Carreiras, Perea, & Grainger, 1997) had participants performing a animal/non-animal semantic decision on words varying for orthographic neighbour. Specifically, the stimuli belonging to the non-animal "category" – that is the stimuli requiring a No response – could have a dense or a sparse orthographic neighbourhood. The authors observed a facilitatory effect of neighbourhood size, that is words with a dense neighbourhood were classified as non-animal faster than words with a sparse neighbourhood.

A words, then, requires more time to be rejected if it has a neighbours belonging to the target category, less time if it has a dense orthographic neighbourhood. This is consistent with what we said earlier in this paragraph: if a word activates its neighbours in the orthographic input lexicon, it also activates their semantic representations, and this influences the performance in semantic tasks. To provide an explanation for those effects we need a semantic module whose architecture is explicitly described, as explicitly described has to be the relations between the semantic module and the orthographic lexicon. Such a module will allow us to explain the above results, and to make new predictions.

6. A semantic module

The semantic module consists of a set of units. Each unit is connected with one unit in the orthographic lexicon (multiple mappings will be discussed in section 8.). We assume that each single unit represents the meaning of the word it is connected with in the lexicon. Connections between the semantic units and the orthographic units are bidirectional and excitatory. Semantic units are organized by category: all the units representing meanings of words belonging to the same category (e.g. Biological Objects) are connected with a unit representing that category. Semantic units and category units are linked by bidirectional excitatory connections. Connections among category units are inhibitory. The system includes a decisional mechanism that monitors the activity of the category units: a words is recognized as belonging to a given category when the activation in the corresponding category unit passes that of the alternative category units by a given amount (criterion).

7. Two explanations and a prediction

Let us first consider how the model incorporating the semantic module can explain the Leotard effect and the orthographic neighbour size effect. *Leotard*. The word *leotard* activates, along with its own lexical representation, the lexical representations of its orthographic neighbours. The lexical representation of *leopard*, then, receives activation. *Leotard* and, although to a smaller extent, *leopard* send activation to the semantic units they are connected with, which send activation to the category units they are connected with. Since both the animal category unit and the non-animal category unit receive activation, the competitions between those category units increases thus delaying the response.

Orthographic neighbourhood size. The density of the neighbourhood was manipulated only for the stimuli not belonging to the category of animals. If also the orthographic neighbours of the stimuli did not belong to the category of animals, then the explanation of the effect would easily follow: when the orthographic neighbourhood is dense, more activity would be sent to the non-animal category unit because more lexical units are activated, causing its activity to grow faster with respect to when the words has only few orthographic neighbours. Therefore, the criterion would be reached faster, and the response made earlier.

Intuitively, it is *unlikely* that stimuli not belonging to the category of animals have neighbours belonging to the category of animals. For example, of the twenty randomly selected Italian words not belonging to the category of animals (*dosso, nastro, monte, mondo, miele, letto, lente, polo, fune, rischio, raggio, pianto, laccio, grotta, freno, alba, anta, vaso, vite, borsa*) with an average neighbourhood size of eight words, only one (*laccio*) turned out having a neighbours (*luccio*) that is the name of an animal; the remaining 159 neighbours were not names of animals. Noteworthy, the situation is reversed if we shift our attention to the animal names. Indeed, it is *likely* that the orthographic neighbour of an animal name do not belong to the category of animals. Therefore, as the number of neighbours of an animal name increase, the number of orthographic neighbours of that name not belonging to the category of animal increases as well.

A prediction. If as the number of orthographic neighbours of an animal name increases, the number of neighbour not belonging to the category of animals increases, the model predicts an effect of orthographic neighbourhood size opposite to that found by Sears et al., for the response "animal" being slower to stimuli with a dense neighbourhood with respect to stimuli with a sparse neighbourhood. The activation sent to the category of animals competes with that sent to the category of non-animals. Thus, as the number of neighbours not belonging to the category of animals increases, the competition increases, and the time taken to reach the criterion increases as well, delaying the response.

8. Some data

To address this issue we designed an experiment (see Mulatti & Job, 2006 for details) where we compared the performance of two groups of participants in two tasks, a free reading and a conditional reading. In the free reading task the participants read all the words they are presented with. As already mentioned, in a reading aloud task the orthographic neighbourhood size exerts a facilitation, that is words with many neighbours are read faster than words with few neighbours. In order to have responses comparable with those of the free reading task, rather than using a semantic decision (that requires a manual response) we decided to use a conditional reading task (Job & Tenconi, 2002). In such task, the participants have to read only the word belonging to a pre-specified category (e.g. animals) and to withhold the response otherwise. Thus, the conditional naming task involves a covert semantic decision, since it is only after having performed a semantic classification that the stimulus can be read, if it belongs to the pre-specified category, or the response withheld, if the stimulus does not belong to the pre-specified category. In our experiment, the participants performing the conditional reading task had to read only the words belonging to the category of Natural Objects.

The predictions are the following: a) in the free reading task, words with many neighbours are read faster than words with few neighbours; b) in the conditional reading task, words with many neighbours are read slower than words with few neighbours.

The material used in the experiment consisted of seventy-two low frequency words (1: note: one of the item was removed from the analyses as nearly half of the participants di not recognized it as a word). Half of them were names of things belonging to the category of Natural Objects (NO), half were names of things belonging to the category of Artefacts (A). In a preliminary test, a pool of participant that did not participate to the main experiment score the typicality of each item as a member of the assigned category. No differences between categories resulted in the analysis of the scores distribution. The experimental items were those of the Natural Objects category. Eighteen experimental items had a dense neighbourhood (mean: 13.4), eighteen experimental items had a sparse neighbourhood (mean: 3.5). Stimuli in the dense and sparse conditions were balanced in terms of typicality, written frequency, and letter length. Noteworthy, the ratio computed comparing the number of neighbours not belonging to the category of Natural Objects with the number of neighbours belonging to the category of Natural Objects was of 5.6 for the dense neighbourhood stimuli, and of 1.8 for the sparse neighbourhood stimuli, $t(33) = 4.2$, $p<.001$.

Fourteen participants performed the free reading task. Each participant was asked to read all the words he was presented with as quickly and accurately as possible. The words appeared in the centre of a computer screen and stayed on until participant responded. The order of presentation of the stimuli was randomized for each participant. The durations of the intervals between the appearance of the stimuli and the onset of the verbal responses constituted the dependent variable (Reaction Times, RTs).

Twelve participants performed the conditional reading task. They were told to read, as quickly and accurately as possible, only the words denoting objects belonging to the category of Natural Objects and to remain silent otherwise. The order of presentation of stimuli was randomized for each participant. The stimuli appeared in the centre of the screen, and stayed on until participant responded. RTs were measured.

Statistical analyses performed on the RTs of correct responses showed that the free reading task was significantly faster than the conditional reading task, and that the main effect of Orthographic Size did not prove significant. Consistently with the prediction, the interaction between the two factors was significant: Whereas in the free reading task words with a dense neighbourhood were read faster than words with a sparse neighbourhood, in the conditional reading task words with a dense neighbourhood were read more slowly than words with a sparse neighbourhood.

The results can be summarized as follow. In the free reading, a task that does not (explicitly) require semantic information to be performed, as the number of orthographic neighbours increases, the time to produce a response decreases: the orthographic neighbourhood size exerts a facilitatory effect.

When the task requires a cover semantic classification, as in the conditional naming task, the characteristics of the semantic representations of the neighbours came into play. Sears et al. (1999) showed that as the number of neighbours belonging to the same category as the target word increases, the process of classifying the word is facilitated. We took this as an evidence for our semantic module: the neighbours send activation to the same category unit as the target word; because of this, the activation in that unit reaches the Criterion for the response faster when the target word has many neighbours. On the other hand, we showed that the increase of the number of neighbours that belong to a category different from that of the target word hinders the semantic classification process. We explained this phenomenon within our framework by postulating that the semantically inconsistent neighbours send activation to a category unit that compete with that of the target word, thus slowing the decision process.

9. Multiple mapping from orthography to semantics, and vice versa

The semantic model we described in section 6. posits that each orthographic unit maps into one single semantic unit, and that each single semantic unit maps into one single orthographic unit. However, Italian (as many other languages including English) counts both ambiguous[5] words, i.e. words that have more than one meaning, and synonyms, i.e. words that have (roughly) the same meaning. An example of the first class of words would be bank (the rising ground bordering a lake or a river; an establishment for the custody of money); examples of the latter would be couch and sofa, which refer to the same thing. These two classes of words pose a problem for the semantic model we have proposed, because such words influence behaviour in idiosyncratic ways and need to be treated as a specific class of words. Specifically, in lexical decision tasks, where participants have to classify strings of letters as word or non-word, ambiguous word are recognized faster than unambiguous words, whereas synonyms are recognized more slowly than non-synonyms (Hino, Lupker & Pexman, 2002). Thus, the so-called ambiguity effect results in a facilitation, the synonymy effect in an interference.

To accommodate for such effects, the way in which the orthographic units are connected with the corresponding semantic units in the model

[5] Psycholinguistics identified two groups of ambiguous words, homonyms words – different meanings – and polysemous words – the meanings that correspond to a polysemous word share a common core meaning. However, Klein and Murphy (2001) showed that even to polysemous words correspond different semantic representations. Because of this, we will include under the same label “ambiguous words” both homonyms and polysemous.

needs to be changed as follows. The orthographic representation of an unambiguous non-synonymic word maps into one single semantic representation. The orthographic representation of an ambiguous word maps into as many semantic representations as the number of meanings that word has. The orthographic representation of a synonyms maps into one single semantic representation, however, this semantic representation maps into as many orthographic representations as the number of synonyms of that word. Such a modified model accounts for both effects.

The ambiguity effect arises because the orthographic representation of an ambiguous word sends activation to more than one semantic units which feed back activation to the orthographic unit. Therefore, since the orthographic unit of an ambiguous word receives activation from many semantic units, its activation grows faster compared to a unit representing an unambiguous non-synonymic word, facilitating its recognition.

The synonymic effect arises because the semantic representation activated by the synonyms sends activation back to the target synonyms but also to the synonyms of the target. Since there is lateral inhibition among units in the orthographic lexicon, the orthographic unit of the non-target synonyms sends inhibition to the orthographic unit of the target synonyms, slowing the raise of its activation, thus hindering its recognition.

10. Concluding Remarks

Implementing a semantic module in the DRC model has allowed us to reach several goals:

(a) to test for the reliability of the model in contexts different from those the model was originally built for;
(b) to define the minimum computational apparatus needed to simulate semantic effects;
(c) to evaluate possible interactions among the already implemented components – namely the orthographic system – and the semantic module;
(d) to test the plausibility of the explanations provided to account for the semantic effects obtained in behavioral experiments.

Above all, the model has been instrumental for deriving predictions, based on its functional architecture and processing assumptions, and to evaluate them against relevant empirical data. Finally, the

and to provide accounts for some phenomena, i.e. ambiguity and synonymy, that

11. References

Norris, D. (2005) How do computational models help us build better theories? In A. Cutler, (Ed.) Twenty-First Century Psycholinguistics: Four Cornerstones.

Lukatela, G., & Turvey, M. T. (1994a). Visual lexical access is initially phonological: 1. evidence from associative priming by words, homophones, and pseudohomophones. Journal of Experimental Psychology: General, 123, 107-128.

Lukatela, G., & Turvey, M. T. (1994b). Visual lexical access is initially phonological: 2. evidence from phonological priming by homophones and pseudohomophones. Journal of Experimental Psychology: General, 123, 331-353.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review, 103, 56-115.

Mulatti, C., & Job, R. (2006). Considerazioni preliminari sull'implementazione di un modulo semantico in un modello computazionale della lettura [Preliminary considerations on the implementation of a semantic module in a computational model of reading]. In P. Giarretta, P. Cherubini, M. Marraffa (Eds.), Cognizione e computazione. CLEUP: Padova.

Pecher, D. (2001). Perception is a two-way junction: Feedback semantics in word recognition. Psychonomic Bulletin and Review, 2001, 8, 545-551.

Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The turple effect. *Memory & Cognition*, 30, 1106-1116.

Rodd, J. (2004). When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. *Psychonomic Bulletin & Review*, 11, 434-439.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. Journal of Experimental Psychology: Human Perception and Performance, 22, 1331-1336.

Broadbent, D. E. (1987). Simple models for experimentable situations. In P.Morris (Ed.), *Modelling cognition* (pp. 169-185). New York: Wiley.

Massaro, D. W. (1992). Understanding mental processes through modeling: Possibilities and limitations. In M. Besson, P. Courrieu, C. Frenck-Mestre, A. M. Jacobs, & J. Pynte (Eds.), *Language perception and comprehension: Multidisciplinary approaches* (pp. 17-18). Marseille, France: Centre National de la Recherche Scientifique, Laboratoire de Neurosciences Cognitives.

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition - Sampling the state of the art. *Journal of Experimental Psychology*, **20**, 1311-1334.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.

Olson, A., & Caramazza, A. (1991). The role of cognitive theory in neuropsychological research. In F. Boller & G. Gratman (Eds.), Handbook of neuropsychology (pp. 287-309). Amsterdam: Elsevier.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Cree, G. S., & McRae, K. (2003), Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, 132, 163-201.

Sartori, G., & Lombardi, L. (2004). Semantic relevance and semantic disorders. *Journal of Cognitive Neuroscience*, 16, 439-452.

Vigliocco, G., Vinson, D. P., & Lewis, W. (2004). Representing the meanings of objects and actions words: the featural and unitari semantic space hypotesis. *Cognitive Psychology*, 48, 422-488.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition*, **15**, 802-814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? Journal of Experimental Psychology: Learning, Memory, and Cognition, **18**, 234-254.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 876-900.

Morton, J. (1961). *Reading, context and the perception of words*. Unpublished PhD thesis, University of Reading, Reading, England.

McClelland, J. L. (1979). On the time relation of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287-330.

Sears, C. R., Lupker, S., & Hino, Y. (1999). Orthographic neighborhood effects in perceptual identification and semantic categorization tasks: A test of the multiple read-out model. *Perception & Psychophysics*, 61, 1537-1554.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 696-713.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of the orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23, 857-871.

Job, R., & Tenconi, E. (2002). Naming pictures at no cost: Asymmetries in picture and word conditional naming. *Psychonomic Bulletin & Review*, 9, 790-794.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming and semantic catgorizationt tasks : Interactions between otrhography, phonology, and semantics. Journal of Experimental Psychology : Learning, Memory, and Cognition, 28, 686-713.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. Journal of Memory and Language, 45, 259-282.