Neuroscience and Biobehavioral Reviews 46 (2014) 85–104

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

Interactive activation and competition models and semantic context: From behavioral to brain data

Markus J. Hofmann[a,b,*], Arthur M. Jacobs[a,c]

[a] Experimental and Neurocognitive Psychology, Free University Berlin, Germany
[b] General and Biological Psychology, University of Wuppertal, Germany
[c] Dahlem Institute for the Neuroimaging of Emotion (D.I.N.E.), Germany

ARTICLE INFO

Article history:
Received 30 August 2013
Received in revised form 14 April 2014
Accepted 23 June 2014
Available online 30 June 2014

Keywords:
Word recognition
Associative spreading
Multiple Read-Out Model (MROM)
Semantic process model
Episodic memory

ABSTRACT

Interactive activation and competition models (IAMs) cannot only account for behavioral data from implicit memory tasks, but also for brain data. We start by a discussion of standards for developing and evaluating cognitive models, followed by example demonstrations. In doing so, we relate IAM representations to word length, sequence, frequency, repetition, and orthographic neighborhood effects in behavioral, electrophysiological, and neuroimaging studies along the ventral visual stream. We then examine to what extent lexical competition can account for anterior cingulate cortex (ACC) activation and the N2/N400 complex. The subsequent section presents the Associative Read-Out Model (AROM), which extends the scope of IAMs by introducing explicit memory and semantic representations. Thereby, it can account for false memories, and familiarity and recollection – explaining why memory signal variances are greater for studied than non-studied items. Since the AROM captures associative spreading across semantic long-term memory, it can also account for different temporal lobe functions, and allows for item-level predictions of the left inferior frontal gyrus' BOLD response. Finally, we use the AROM to examine whether semantic cohesiveness can account for effects previously ascribed to affective word features, i.e. emotional valence, and show that this is the case for positive, but not for negative valence.

Contents

* Corresponding author at: Department of Psychology, General and Biological Psychology, Room Z.01.11, Max-Horkheimer Str. 20, 42119 Wuppertal, Germany. Tel.: +49 202 4392340; fax: +49 202 439 2926.
E-mail address: mhofmann@uni-wuppertal.de (M.J. Hofmann).

http://dx.doi.org/10.1016/j.neubiorev.2014.06.011
0149-7634/

1. Introduction

Back in 1981, the interactive activation and competition model (IAM) was a major step ahead in cognitive modeling for several reasons. Sternberg's (1969) seminal model of verbal working memory, or Morton's (1969) logogen model already had zoomed into the blackbox between stimulus and response, breaking it up into specific serial or parallel stages of information processing (i.e., the famous boxes and arrows of 'boxological models'; cf. Jacobs and Grainger, 1994). The IAM combined features of previous formal word recognition models by Broadbent (1967), Morton (1969), Rumelhart and Siple (1974), or Treisman (1978) with pioneering "neural models" of the time (e.g., Anderson et al., 1977; Grossberg, 1980). It was the first model that really zoomed into (the dynamics of) those 'boxes' and allowed to simulate the time course of information processing in several parallel layers (i.e., feature, letter, and word unit layer; Fig. 1).

The IAM also implemented two neurally plausible features, connectivity (excitation and inhibition allowing within-level competition) and interactivity (top–down feedback allowing between-level memory effects on perceptual processing). Both features were disputed at the theoretical level by main stream cognitive modelers favoring modular cognitive architectures at the time (e.g., Massaro, 1988; Massaro and Cohen, 1991; Paap et al., 1982), but they were also experimentally testable (Jacobs and Grainger, 1992; McClelland, 1991). By making the top–down feedback algorithmically concrete, the IAM succeeded in elegantly simulating the word superiority effect (Fig. 2). This corresponded to Helmholtz's idea of *unconscious inferences*, expressing his belief that sensory data are modified by previous experience via ideas/concepts, before they become a true perception (Boring, 1950; Grossberg, 1980). In the more modern words of Grossberg (1980) "sensory data activate a feedback process whereby a learned template, or expectancy, deforms the sensory data until a consensus is reached between what the data are and what we expect them to be. Only then do we perceive anything". In Friston's (2010) unifying principle of brain function, such feedback processes (mathematically formulated within the frameworks of free energy and predictive coding) play a central role and there is now ample evidence for its neural plausibility (e.g., Price and Devlin, 2011).

The IAM was perhaps the first model in this field that made all information processing steps between input and output fully transparent, thus providing a comprehensive description of information processing at the micro level, and – achieving what can be considered a gold standard of model evaluation criteria (Jacobs and Grainger, 1994) – it predicted a new phenomenon which had previously not been observed: the neighborhood frequency effect (cf. Jacobs et al., 1998). This effect was experimentally confirmed (Grainger et al., 1989) and thereby fired further developments leading to many successful extensions or variants of the basic interactive activation architecture, e.g. the model of the Stroop task (Cohen et al., 1990), the Dual Read-Out Model (DROM; Grainger and Jacobs, 1994), the Multiple Read-Out Model (MROM; Grainger and Jacobs, 1996) and its extension including phonological processing units (MROM-p; Jacobs et al., 1998), the conflict monitoring theory (CMT; Botvinick et al., 2001), the dual-route cascaded model (DRC, Coltheart et al., 2001), the connectionist dual-process model (CDP++; Perry et al., 2007, 2010), or the recent AROM including an implemented semantic layer (Hofmann et al., 2011), to name only a few.

While during the 80s and 90s IAMs were very successful in predicting behavioral data such as error rates, or response time means and distributions in many different tasks, in 1995 only Jacobs and Carr (1995) speculated how they could be applied to neuroimaging data, and how functional neuroimaging could be used to constrain computational models of cognition in general: (1) by providing information about the neuroanatomical loci of different subprocesses and hence system decomposability and (2) by delineating the temporal dynamics of the cognitive process(es) under investigation (cf. Barber and Kutas, 2007). It took a few years until

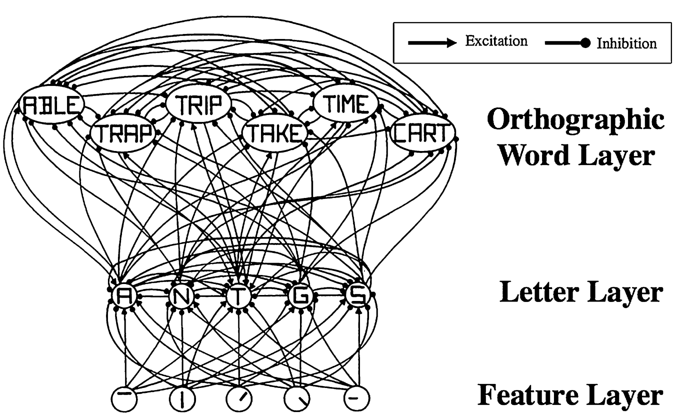

Fig. 1. Architecture of the classic IAM. For each letter position, there are visual feature units in a feature layer. For instance, if a "T" is presented to the model at the first position, the visual features "|" and "–" activate the unit "T" at the letter layer, which in turn activates all units at the orthographic word layer starting with a T, e.g. trip or take.
Adopted from McClelland and Rumelhart (1981).

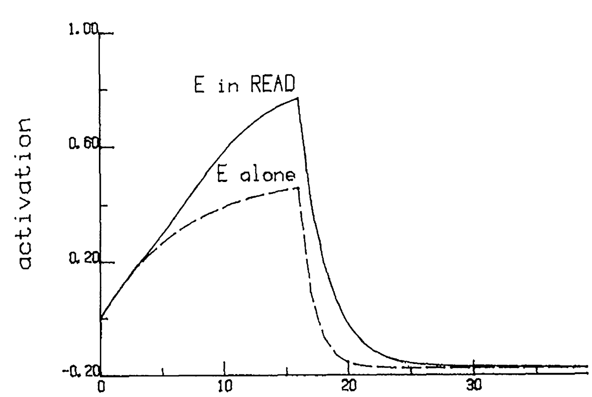

Fig. 2. Example simulations that account for the word superiority effect. Perceptual identification of a letter is faster, if it is contained in a word. The classic IAM can account for this by the letter level activations shown at the *y*-axis. The *x*-axis represents model cycles. When the identified target letter obtains excitation from the orthographic word unit 'READ', its activation becomes greater than when the letter is presented in isolation. While greater activations indicate greater evidence that the letter has been presented, the IAM accounts for a faster and less error-prone identification of letters in words.
Adopted from McClelland and Rumelhart (1981).