# Language Processing, Computational Representational Theory of Mind and Embodiment: Inferences on Verbs

**J. Ezquerro and M. Iza**

**Abstract** The computational representational theory of mind conceptualized by Jerry Fodor has dominated the schedule of cognitive science for many years. Fodor defended it as an empirical theory whose plausibility depends on facts about how the mind works. However, current developments in philosophy, psycholinguistics and neurosciences present a challenge to this received view. These developments assume embodied and situated cognition. This paper reviews and assesses the prospects of these new accounts, focusing on processing verbs and on the role of emotions in language processing.

## 1 Introduction

Sometimes, philosophical theses fix the schedule of empirical research. This is the case of Jerry Fodor. In *The Language of Thought* (1975), Fodor defended the computational representational theory of mind (CRTM) and through the years developed the main conceptual consequences of such a view. Two of these consequences were methodological individualism and methodological solipsism. In *Psychosemantics* (1987), Fodor characterizes these theses in the following way:

> Methodological individualism is the doctrine that psychological states are individuated *with respect to their causal powers*. Methodological solipsism is the doctrine that psychological states are individuated *without respect to their semantic evaluation* (p. 42)

J. Ezquerro (✉)
Institute for Logic Cognition Language and Information (LCLI),
University of the Basque Country, UPV/EHU, Donostia-San Sebastián, Spain
e-mail: jesus.ezquerro@ehu.eus

M. Iza
Department of Psychology, University of Malaga, Málaga, Spain
e-mail: iza@uma.es

These two theses constitute the core of what is usually known as "the received view". They go hand in hand, although Fodor tried to distinguish them carefully. Methodological individualism opposes anti-individualism defended by Tyler Burge:

> Anti-individualism is the view that not all of an individual's mental states and events can be type-individuated independently of the nature of the entities in the individual's environment. There is, on this view, a deep individuative relation between the individual's being in mental states of certain kinds and the nature of the individual's physical or social environments...... ([9], p. 46)

It is clear that this characterization appeals to relational properties (physical and social environments) in order to identify psychological states, and Fodor concedes that it could be so *if and only if* these relational properties can affect causal powers. Thus, Fodor thinks that individualism is a conceptual point: if something lacks causal powers, then it lacks of explanatory force. Therefore, it should be ruled out from psychological explanation.

However, at this point methodological solipsism enters the scene. Far from being a purely conceptual point, Fodor qualifies it as an empirical thesis:

> 'Methodological solipsism' is, in fact, an empirical theory about the mind: it's the theory that mental processes are computational, hence syntactic. I think this theory is defensible; in fact, I think it's true. But this defense can't be conducted on a priori grounds, and its truth depends simply on the facts about how mind works. (1987, p. 43)

In this way, methodological solipsism constrains and sustains methodological individualism: even if the later does not prohibit relational properties, the former does. Given that semantical properties are relational, they fall out of the scope of scientific explanation [23]. The story is well known. Realizing that we cannot do without semantics, Fodor engaged with a program—Psychosemantics—trying to account for meaning in a way that respects the fundamentals of CRTM. However, this program was abandoned because it was unable to respond to some conceptual challenges, especially, twin earth cases.

Since then, time has shown that the computational representational theory of mind is probably false, at least in the way J. Fodor characterized it. On the philosophical side, Burge's anti-individualist position, and A. Clark's extended mind thesis go in the opposite direction and have gained acceptance [10, 12]; Furthermore, inasmuch as Fodor considers methodological solipsism as an empirical theory, our purpose is to revise and assess the current developments in cognitive psychology and cognitive linguistics in order to show that the direction of research has clearly departed from the received view, and its results and findings go against it. Our attention will concentrate on language processing, and within this field, on the case of verbs.

Current research shows that visual, motor and emotional information are active during sentence processing [1, 24, 43]. For this reason, embodied theories of cognition reject the received view. This view affirms that meaning depends only on the structural relations with other symbols, following the assumption of methodological solipsism above mentioned. Embodied accounts suggest, instead, that symbols have to be grounded to their referents in the environment [36, 63].

There are two hypotheses concerning the storage of meaning in cognitive psychology. The received view defends the complexity hypothesis, assumed by Fodor [21], Kinstch [46] and Thorndyke [73]. This hypothesis affirms that a word with many semantic components would require more processing resources, comprehension time and long-term memory space than a word with a smaller number of components, thus interfering more strongly with memory for surrounding words.

The connectivity hypothesis, on the contrary, considers verb semantic structures as frames for sentence representation, claiming that memory strength between two nouns in a sentence increases with the number of underlying verb subpredicates that connect the nouns.

Following the above assumptions, therefore, the complexity hypothesis predicts that a verb with many subpredicates will lead to poorer memory strength between the surrounding nouns than a verb with few subpredicates. The connectivity hypothesis, instead, predicts that verbs with many subpredicates will lead to "greater" memory strength between nouns in cases where the additional subpredicates provide semantic connection between the nouns.

The complexity hypothesis assumes the "bin" view of memory, in which the capacity limitations of various stages of memorial processing form a central theoretical notion. The connectivity hypothesis, instead, holds a structural view of memory, in which the representational assumptions are crucial.

Gentner [25] tested the memory predictions of the above two accounts. In three experiments, subjects recalled subject–verb–object sentences, given subject nouns as cues. General verbs with relatively few subpredicates were compared with more specific verbs whose additional subpredicates either did or did not provide additional connections between the surrounding nouns. The level of recall of the object noun, given the subject noun as cue, was predicted by the relative number of "connecting" subpredicates in the verb, but not by the relative "number" of subpredicates. Results did support the connectivity hypothesis over the complexity hypothesis. Gentner interpreted the results in terms of a model in which the verb conveys a structured set of subpredicates that provides a connective framework for sentence memory.

The proposed model, called the central components model, differs from both the extreme decompositional model and the extreme meaning-postulate model. Here, the verb's representation is aimed to specify the pattern of inferences that is most dependably activated when the verb is comprehended. The model is clearly decompositional: it assumes that one verb leads to several separable (though structurally related) inferences, and that both lexical generalities and psychological phenomena can be stated in terms of connected sets of subpredicates embodying these inferences.

However, these inferences are not intended to embody necessary-and-sufficient conditions for use. Instead, the representation offered for a given verb expresses the central set of inferences (the set most frequently and reliably associated with the verb's use). The representations are not considered as exhaustive; indeed, it is very clear that they are not. For instance, the verb "give" clearly has other possible inferences: that the giver is generous, that she has the means to give away objects

and so on. There is no fixed stopping point for this kind of inferential processing. Furthermore, the subpredicates are not required to be atoms belonging to a primitive base. A component is useful in a psychological representation if it functions as a familiar unit at that level of representation. Components at one level of representation may be decomposed at the next level down into a further network of linked components.

These results are compatible with the assumption that inferencing begins on-line, before the syntactic parser is completed. Inferencing would begin with the central set and continue, radiating outward more esoteric inferences, if the central set of inferences is not sufficient for a satisfactory interpretation of the sentence [14], or, even, if nonlinguistic cues provide enough information to spread forward inferences [40].

The importance of perceptual and action simulations has been recurrently stressed in the area of concrete language processing. Researchers emphasize the importance of embodiment by either amplifying evidence for the activation of embodied representations, or by establishing boundary conditions (such as temporal overlap and integratability in the domain of perception, linguistic focus, grammar, affordances and the kind of effectors involved in response to the domain of action).

For example, Glenberg and Kaschak [27] asked participants to view series of sensible sentences that either described transfer toward a reader ("Open the drawer") or away from the reader ("Close the drawer"), as well as series of nonsense sentences ("Boil the air") that did not entail any transfer. The task was to judge whether sentences made sense by pressing one of the vertical keys on a three-button box that required a movement either towards or away from the body. They found that responses were faster when the motion implied by the sentence matched the actual hand motion (action sentence compatibility effect).

In a similar way, Zwaan and Taylor [78] showed that this effect is quite specific: an activation of compatible motor responses was localized on the verb region of the sentence, but not on the preverb, postverb and sentence-final regions. Then, it was tested whether maintaining focus on the action, by following the verb with an adverb implying action, might cause motor resonance to affect both the verb and the adverb that follows it (focus on the action, "When he saw a gas station, he exited slowly", versus on the agent, "When he saw a gas station, he exited eagerly"; see also [70].

From now on, we will concentrate on three issues. First, our review focuses on sentence and discourse processing, whereas most of the previous reviews focus on single-word processing [45]. Our previous research stressed the importance of discussing the role of sensory-motor and affective processes in comprehension of language segments that provide enough context to constrain the interpretation of verb meaning [18, 20, 38]. Second, we will emphasize the trajectory of embodiment research in the domain of verb inferencing, attending to behavioral and neuroscientific data. In our opinion, these data put in jeopardy the received view [13, 71, 76, 8]. Finally, we will discuss a variety of old and new theoretical approaches to understanding of the role of inferences and verbs in language processing.

## 2   Automatic Processing

Grammatical categories across languages distinguish between nouns and verbs. This distinction tend to rely on lexical-semantic criteria (nouns refer to entities, while verbs refer to actions or events), and syntactic or distributional criteria (nouns and verbs play different roles in sentences and occur in conjunction with different sets of grammatical morphemes). In this sense, all languages have a mechanism for mapping lexical items with particular semantic structures to specific syntactic roles, which in some languages are marked morphologically (e.g. Basque language). These different criteria are engaged specifically in the lexical-semantic processing of verbs and nouns, as semantic priming studies show.

Semantic priming refers to the processing advantage that occurs when a word (the target) is preceded by another word (the prime) when it is related in meaning to the target. The nature of the meaning relationship can be due to shared physical, functional or visual features and membership in the same semantic category. The phenomenon of priming has become ubiquitous within cognitive psychology, and the existence of positive semantic priming has been solidly established [7, 69].

The automatic character of this process is shown by experimental data from a lexical decision task (see, e.g. [48]). In these experiments, subjects were asked to decide as quickly as possible whether each occurrence of a letter string was an English word, and to respond in order to provide reaction times. The left side of the screen was taken up by a patch containing computer-related verbal "garbage", which subjects were instructed to ignore, but sometimes contained a word planted to relate to one of a subject's current concerns. When the target string was indeed a word, the reaction time of reporting this was significantly slower if the distractor patch contained a concern-related word. Thus, concern-related stimuli seem to impose an extra load on cognitive processing even when they are peripheral and subjects are consciously ignoring them. This finding adds a further face to the automaticity of the effect (see also [44]).

Besides, this effect was shown in another cognitive process with a modified Stroop procedure [64]. For instance, MacKay et al. [55] demonstrated three taboo Stroop effects that occur when people name the color of taboo words (e.g. death, war). These effects are (i) longer color-naming times for taboo than for neutral words, an effect that diminishes with word repetition; (ii) superior recall of taboo words in surprise memory tests following color naming; (iii) better recognition memory for colors consistently associated with taboo words rather than with neutral words. They argue that taboo words trigger specific emotional reactions that facilitate the binding of taboo word meaning to salient contextual aspects.

Emotion Stroop experiments found little evidence of automatic vigilance, for instance, slower lexical decision time or naming speed for negative words after controlling for lexical features. Estes and Adelman [17] analyze a set of words, controlling for important lexical features, and find a small but significant effect for word negativity and conclude that this effect is categorical. Larsen et al. [50] analyze the same data set but include the arousal value of each word. They find

non-linear and interaction effects in predicting lexical decision time and naming speed. Not all negative words produce the generic slowdown. Only negative words that are moderate to low on arousal produce more lexical decision time slowing than negative words higher on arousal. Similarly, Kahan and Hely [41] showed that the roles of valence and word frequency interact in contributing to the emotional Stroop effect.

Altarriba and Canary [4] also examined the activation of arousal components for emotion-laden words in English (for instance, kiss, death) in two groups of monolingual (English) and bilingual (Spanish–English) subjects. Prime–target word pairs were presented for lexical decisions to English word targets in either high arousal, moderate or unrelated conditions. Results revealed positive priming effects in both arousal conditions for both groups of subjects. Nevertheless, while the baseline conditions were similar across groups, the arousal conditions produced longer latencies for bilinguals than for monolinguals.

## 3   The Assumption in Cognitive Linguistics Regarding Verb Taxonomy May Be in Error

A widely shared assumption in cognitive linguistics [49, 39] is that verb representations are organized in a taxonomic structure during cognition processes. Also, it is considered that this taxonomy resembles the one of noun prototype structure and taxonomy found by Rosch [66–68]. There is however little or no empirical evidence supporting the hypothesis that verbs are associated through taxonomic structures.

A cross-linguistic (Finland–Spain) study on categorization of lexemes indeed suggests that the psycholinguistic processes relative to the categorization of verbs lack taxonomic structures [37]. The subjects (100 subjects in each country) were unable to associate either the subordinate or the superordinate verbs suggested in cognitive linguistics [49, 39]. The task was formulated in four different wordings for four different groups as well as controlled by tasks using nouns. The subjects performed according to the taxonomy hypothesis on nouns but not on verbs.

A small-scale developmental study with teenagers aged between 13 and 19 years showed that the ability to associate in taxonomic structures with nouns was age related. The younger group made more errors, while the age group 18–19 performed successfully. This suggests that there are questions requiring further study, for example, whether the taxonomy-principle is an inherited property of the mental lexicon or a relatively late-learned association pattern.

In dialogue situations, it is usual to assume that partners know lots of things, so that agents in communication make explicit only those facts they believe that the partner does not know. This shared implicit knowledge is also known as common-sense knowledge. Due to this fact, it is difficult or impossible to analyze a discourse without

relying on a large background of knowledge. For instance, this is the case of some computer programs that try to establish a relationship between logic and language. In neurolinguistic programming (NLP) systems, often verbs are considered as predicates and verb valences as arguments. In this approach, new sentences or propositions can be inferred from the discourse. Furthermore, an inference system for verb frames is used and an evaluation process is normally required.

Learning inference relations between verbs and propositions is at the heart of many semantic applications. However, most prior work on learning such rules was focused on a narrow set of information sources, that is to say, mainly distributional similarity and to some extent manually constructed verb co-occurrence patterns (see, e.g. [6]). In this regard, we claim below that it is imperative to use information from various discursive scopes to provide: a much richer set of inferential cues to detect entailment between verbs and combine them as features in an unsupervised classification framework.

## 4 Self-organizing Maps

Within the emergentist paradigm [16, 56], recent studies [52, 51] interpret the interaction among the categories of "aspect" and "actionality" as the result of the learners' analysis of the probabilities of co-occurrences between aspect morphology and actionality semantics in the linguistic input. Children extract from the input the statistical frequencies of the combinations between aspect forms and actionality classes. Initially, they strengthen the production of the most frequent associations, until prolonged exposure to the input and the increasing amount of data from the input reduce the statistical difference between the most and least frequent combinations.

Li and Shirai [51] tested this hypothesis by means of computational simulations based on self-organizing maps [47]. These maps consist in supervised associative neural networks of "knot receptors" that classify input data, translating relationships of similarity into topological relationships of proximity. Through incremental exposure to an increasing amount of data, the receptors are topologically organized on the network in such a way that associated receptors have the tendency to recognize homogeneous sets of data. These maps are biologically plausible models: human cerebral cortex can be conceived as essentially a multiple feature-map, where all neurons are initially co-activated and the associative strengths between neurons become more focused in parallel with distributional increase of the corresponding co-occurrences in the input.

If the acquisition of the "semantic core" of a construction (and of the semantics of particular actions) and entrenchment (verb frequency) are both probabilistic and incremental processes, then it is easier to imagine how the two processes could work together. For example, perhaps the frequency effect observed is mediated by the effect of verb semantics. Each time a verb is listen in a new context, the learner is given a further opportunity to increase her knowledge of the precise semantics of

the verb, and hence the constructions in which the verb can appear. Each time the learner finds a causative construction (and infers the speaker's intended meaning), this constitutes evidence that the construction encodes direct external causation. We think that this result fits very well with the general account proposed by Clark [11] that conceives the brain as a predictive system.

Alishahi and Stevenson [3] present a computer model that makes use of this statistical correlation between semantics and syntax. The model takes frames that are representations of scene–utterance pairings as input (i.e. of the utterance heard by the child and the scene it describes). These representations specify the verb (e.g. take), its semantic primitives (cause, move), argument roles (agent, theme goal), categories (human, concrete, destination-predicate) and syntactic structure (arg1 verb arg2 arg3). Whenever the model encounters a new frame, it stores the semantic and syntactic information in the lexical entry for the verb. It then groups this frame together with the existing class of frames whose members share the greatest number of syntactic and semantic features, or, if none are sufficiently similar, creates a new class (since the classes are not prespecified, this will be the case for the majority of frames in the early stages of learning). Whenever the model is presented with a previously encountered frame, it increases the representational strength of that frame and of its parent class.

To simulate generalization, every five training pairs, the model is presented with a frame from which the syntactic representation has been deleted, but with all other information intact. The task of the model is to choose the most appropriate syntactic structure. In order to choose the most appropriate pattern, the simulation can make use of two sources of information. These sources are, on the one hand, the syntactic structures stored in the lexical entry for that particular verb (item-based knowledge), and, on the other, those stored in the lexical entry of the other verbs in the same class (i.e. the other verbs to which this verb is semantically and syntactically similar; class-based knowledge).

## 5   Pragmatics and Emotional Inferences

As happens in dialogue contexts, it is usually taken for granted that, in the course of reading a text, world knowledge is often required in order to establish coherent links between sentences [58, 77]. The content grasped from a text turns out to be strongly dependent upon the reader's additional knowledge that allows a coherent interpretation of the text as a whole (see [31]).

The world knowledge directing the inference may be of a distinctive nature. Gygax et al. [33] showed that mental models related to human action may be of a perceptual nature and may include behavioral as well as emotional elements. However, Gygax [32] showed the unspecific nature of emotional inferences and the prevalence of behavioral elements in readers' mental models of emotions. Inferences are performed in both directions; emotional inferences are based on behavior, and vice versa.

Harris et al. [15] and Pons et al. [62] showed that different linguistic skills—in particular lexicon, syntax and semantics—are closely related to emotion understanding. Iza and Konstenius [37] showed that additional knowledge about social norms affects the predictions of participants about what should be inferred as the behavioral or emotional outcome of a given social situation.

Syntactic and lexical abilities are the best predictors of emotion understanding, but making inferences is the only significant predictor of the most complex components (reflective dimension) of emotion comprehension in normal children. Recently, Farina et al. [19] showed in a study that the relation between pragmatics and emotional inferences might not be so straightforward. Children with high functioning autism and Asperger syndrome present similar diagnostic profiles, characterized by satisfactory cognitive development, good phonological, syntactic and semantic competences, but poor pragmatic skills and socio-emotional competences. After training in pragmatics, descriptive analyses showed that the whole group displayed a deficit in emotion comprehension, but high levels of pragmatic competences. Furthermore, this indicates the necessity of studying the relationship between emotion and inference in normal subjects too.

Vanhatalo [75] showed that a group of synonyms of speech act verbs actually had semantically distinctive emotional elements as well as different social norms associated to these lexemes. This semantic knowledge related to inferences was not present either in dictionaries or in current literature, increasing demands for empirical studies directed at native speaker intuitions.

We also suggest that, while behavioral elements may indeed be of perceptual nature and the inference between emotion and behavior less culturally dependent, especially when concerned with basic emotions the inference concerned with social norms may be more complex, requiring elaborative inference. We suggest that, in further studies, a distinction between basic emotions and nonbasic emotions, social settings and non-social settings should be taken into account. The cognitive models concerned with social action may be of a more complex nature, but with recognizable features on lexical and syntactic levels.

An increasing amount of research has been done with the aim of testing whether reenactment of congruent and incongruent emotional states affects language comprehension. For example, Havas et al. [34] asked subjects to read sentences describing emotional or nonemotional events while being in a matching or mismatching emotional state. The major result was that sentences describing pleasant events were processed faster when subjects were smiling (pen in the teeth condition). Contrarily, unpleasant sentences were processed faster when subjects were prevented from smile (pen in the lips condition). Later, Havas et al. [35] injected subjects with botulinum toxin A to temporarily paralyze a facial muscle responsible for frowning. Subjects were instructed to read sad and angry sentences. They found that the reading of sad and angry sentences was slowed after Botox injections, suggesting that being prevented from frowning makes it more difficult to simulate sadness and anger.

Using a different procedure, namely electromyography (EMG) measurement of the zygomatic major and corrugator supercilii muscle regions, Foroni and Semin

[24] found that motor resonance was induced when participants processed adjectives describing emotion (happy, sad), but to a lesser extent than when participants processed action emotional verbs (to smile, to frown). Finally, Lugli et al. [54] asked subjects to judge the sensibility of positively and negatively valenced sentences describing toward transfer ("The object is nice. Bring it toward you".) and away transfer ("The object is ugly. Give it to another person".) while responding with a mouse in a direction that either matched or mismatched the direction described by the sentence. They found that responding to sentences describing positive objects was faster with a "toward the body" hand motion. Likewise, responses to sentences describing negative objects were faster with an "away from the body" hand motion.

To sum up, the results of all these studies suggest that sentence comprehension is facilitated when the suggested mood of the sentence is congruent with the concurrent mood of the comprehender.

## 6   Discussion

Successful language learning combines generalization and the acquisition of lexical constraints. This conflict is especially clear for verb argument structures, which may generalize to new verbs or resist generalization with certain lexical items. In this sense, we have to capture the emergence of feature biases in word learning. In this work, we have described a model capable of learning about structure variability simultaneously on several levels. This architecture makes inferences about construction variation and applies them when faced with verbs for which it has very little data. Even more, this framework can try to explain the slight divergences between model predictions and human behavior (see, e.g. [28, 63]).

The model action-based language [26] offers an action-based account of language comprehension by making use of neurophysiologic findings on mirror neurons [60], and by adopting controller and predictor models from theories of motor control responsible for computing goal-oriented motor commands and predicting sensori-motor effects of these commands [30]. Here, language comprehension is tantamount to predicting sensorimotor and affective effects of the performed action. For instance, upon hearing the word "walk", a person's speech mirror neurons activate an associated action controller responsible for generating the motor commands necessary for interaction. Then, the predictor (sensory, motor or emotional) of the target word establishes possible consequences of the action to be performed; That is to say, under this approach, the same hierarchical mechanisms that are used in controlling action (control and predictor) are utilized for generating grammatical consequences in language processing.

Although it remains limited in scope, the model presented here lays the basis for an approach to entailment detection that combines a robust semantic representation with inference performance. We have illustrated the potential of the approach by showing how it could handle a limited range of interaction between nominal

predication, verbal predication and cognitive inferences. Current work concentrates on extending the model coverage and on evaluating it on a full-size benchmark designed to illustrate a wider range of focused inferences (such as affective or status relations) between basic predications and the various semantic phenomena present in the discourse (see, e.g. [68]).

Several studies have been concerned with the brain regions exhibiting different responses to noun and verb processing tasks. Such differences manifest the location of verb-specific processing and the activity patterns within common verb-specific regions.

Evidence from aphasic patients suggests that noun deficits tend to arise from lesions to the left middle-to-anterior temporal lobe. Verb deficits resulted from lesions in left inferior frontal regions, premotor regions, or in posterior temporal to inferior parietal regions [2]. Many imaging studies have found greater activation for verbs compared with nouns in the left prefrontal cortex [5, 59] and in left posterior temporal regions [53, 74].

Recently, Willms et al. [76] showed greater activity for verbs than for nouns in Spanish–English bilinguals in four regions: left posterior middle temporal gyrus, left middle frontal gyrus, pre-supplementary motor area and right middle occipital gyrus. Their results suggest that the neural substrates underlying verb-specific processing are largely independent of language in bilinguals.

When people become committed to pursue a goal, that event initiates an internal state termed a "current concern". One of the properties of this state is potentiating emotional reactivity to cues associated with the goal pursued. The emotional responses thus emitted begin within about 300 ms after exposure to the cue—early enough to be considered purely central, nonconscious responses at this stage. Because they appear to be incipient emotional responses but lack many of the properties normally associated with emotion, they are called "protoemotional" responses (see also [57]).

These responses go on in parallel with early perceptual and cognitive processing, with which they trade reciprocal influences. The intensity (and other features) of protoemotional responses affects the probability that the stimulus will continue to be processed cognitively. The results of continued cognitive processing in turn modulate the intensity and character of the emerging emotional response. That goal setting affects performance at different phases of skill acquisition and for different levels of ability [42]. Furthermore, there are different distributions of cognitive processing and mental content in different phases of goal-striving sequences [29].

As these experimental data show, activating accessible constructs or attitudes through one set of stimuli can facilitate cognitive processing of other stimuli under certain circumstances, and can interfere with it under other circumstances. Some of the results support and converge on those centered on the constructs of current concern and emotional arousal (see [72]).

Future research should take seriously into account the following question: how to develop models where emotion interacts with cognitive processing. One example could be the work of Pitterman et al. [61] where speech-based emotion recognition is combined with adaptive human–computer modeling. With the robust recognition

of emotions from speech signals as their goal, the authors analyze the effectiveness of using a plain emotion recognizer, a speech–emotion recognizer combining speech and emotion recognition, and multiple speech–emotion recognizers at the same time. The semistochastic dialogue model employed relates user emotion management to the corresponding dialogue interaction history and allows the device to adapt itself to the context, including altering the stylistic realization of its speech.

Summing up, both the new directions of research in cognitive linguistics and the empirical results of brain sciences indicate that the received view about how the mind works should be abandoned. Special emphasis must be put on the role of emotions in explaining cognitive processing, given the importance they have in order to explain language comprehension and performance. But this can only be done assuming embodied and situated cognition.

# References

1. Borghi, A.M., Di Ferdinando, A., Parisi, D.: Objects, spatial compatibility, and affordances: a connectionist. Study. Cogn. Syst. Res. **12**, 33–44 (2011)
2. Aggujaro, S., Crepaldi, D., Pistarini, C., Taricco, M., Luzzatti, C.: Neuro-anatomical correlates of impaired retrieval of verbs and nouns: interaction of grammatical class, imageability and actionality. J. Neurolinguist. **19**, 174–194 (2006)
3. Alishahi, A., Stevenson, S.: A cognitive model for the representation and acquisition of verb selectional preferences. Proceedings of the ACL-2007 Workshop on Cognitive Aspects of Computational Language Acquisition, pp. 41–48. Czech Republic, Prague (2007)
4. Altarriba, J., Canary, T.M.: The influence of emotional arousal on affective priming in monolingual and bilingual speakers. J. Multiling. Multicult. Dev. **25**(2), 248–265 (2004)
5. Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., Saxe, R.: Concepts are more than percepts: the case of action verbs. J. Neurosci. **28**, 11347–11353 (2008)
6. Berant, J., Dagan, I., Goldberger, J.: Learning entailment relations by global graph structure optimization. Comput. Linguist. **38**(1), 73–111 (2012)
7. Besner, D., Stolz, J.A.: Unintentional reading: Can phonological computation be controlled? Can. J. Exp. Psychol. **52**, 35–42 (1998)
8. Binder, J.R., Desai, R.H.: The Neurobiology of semantic memory. Trends. Cogn. Sci. **15**(11), 527–536 (2011)
9. Burge, T.: Philosophy of language and mind: 1950–1990. Philos. Rev. **101**, 3–52 (1992)
10. Clark, A.: Language, embodiment and the cognitive niche. Trends Cogn. Sci. **10**(8), 370–374 (2006)
11. Clark, A.: Whatever next? Predictive brains, situated agents, and the future of Cognitive Science. Behav. Brain Sci. **36**, 181–253 (2013)
12. Clark, A., Chalmers, D.: The Extended Mind. Analysis **58**, 7–19 (1998)
13. Coello, Y., Bartolo, A. (eds.): Language and Action in Cognitive Neuroscience. Psychology Press, New York (2012)
14. Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychol. Rev. **82**, 407¬–428 (1975)
15. De Rosnay, M., Harris, P.L.: Individual differences in children's understanding of emotion: the roles of attachment and language. Attach. Hum. Dev. Apr 4(1), 39–54 (2002)

16. Ellis, N.C.: Constructions, chunking and connectionism: the emergence of second language structures. In: Doughty, C.J., Long, M.H. (eds.) The Handbook of Second Language Acquisition, pp. 63–103. Blackwell, Oxford (2003)
17. Estes, Z., Adelman, J.S.: Automatic vigilance for negative words in lexical decision and naming. Emotion **8**(4), 441–444 (2008)
18. Ezquerro, J., Iza, M.: Modeling the interaction of emotion and cognition. Int. J. Res. Rev. Appl. Sci. **17**(3) (2013)
19. Farina, E., Albanese, O., Pons, F.: Making inferences and comprehension of emotions in children of 5–7 years of age. Psychol. Lang. Commun. **11**(2), 3–19 (2007)
20. Fischer, M.H., Zwaan, R.A.: Grounding cognition in perception and action. Q. J. Exp. Psychol. Sect. A—Hum. Exp. **61**, 825–857 (2008)
21. Fodor, J.: The Language of Thought. Harvard University Press, Cambridge, Mass (1975)
22. Fodor, J.: Psychosemantics. The Problem of Meaning in the Philosophy of Mind. MIT Press, Cambridge, Mass (1987)
23. Fodor, J.: The Mind Doesn't Work That Way. The Scope and Limits of Computational Psychology. MIT Press, Cambridge, Mass (2000)
24. Foroni, F., Semin, G.R.: Language that puts you in touch with your bodily feelings: the multimodal responsiveness of affective expressions. Psychol. Sci. **20**(8), 974–980 (2009)
25. Gentner, D.: Verb semantic structures in memory for sentences: evidence for componential representation. Cogn. Psychol. **13**, 56–83 (1981)
26. Glenberg, A.M., Gallese, V.: Action-based language: a theory of language acquisition, comprehension and production. Cortex **48**, 905–922 (2012)
27. Glenberg, A.M., Kaschak, M.P.: Grounding language in action. Psychon. Bull. Rev. **9**, 558–565 (2002)
28. Glenberg, A.M., Webster, B.J., Mouilso, E., Havas, E., Lindeman, L.M.: Gender, emotion and the embodiment of language comprehension. Emot. Rev. **1**, 151–161 (2009)
29. Gollwitzer, P.M.: The volitional benefits of planning. In: Gollwizer, P.M., Bargh, J.A. (eds.) The Psychology of Action: Linking Cognition and Motivation to Behavior, pp. 287–312. Guilford Press, NY (1996)
30. Grush, R.: The emulation theory of representation: motor control, imagery and perception. Behav. Brain Sci. **27**, 377–442 (2004)
31. Guendouzi, J., Loncke, F., Williams, M.J. (eds.): The Handbook of Psycholinguistic and Cognitive Processes. Psychology Press, New York (2011)
32. Gygax, P.M.: L'inférence émotionnelle durant la lecture et sa composante comportementale. [Emotion inference during reading and its behavioral component]. L'Année psychologique **110**, 253–273 (2010)
33. Gygax, P., Tapiero, I., Carruzzo, E.: Emotion inferences during reading comprehension: what evidence can the self-pace reading paradigm provide? Discourse Process. **44**, 33–50 (2007)
34. Havas, D.A., Glernberg, A.M., Rinck, M.: Emotion simulation during language comprehension. Psychon. Bull. Rev. **14**, 436–441 (2007)
35. Havas, D., Glenberg, A., Gutowski, K., et al.: Cosmetic use of botulinum toxin-A affects processing of emotional language. Psychol. Sci. **21**(7), 895–900 (2010)
36. Horchak, O.V., Giger, J.C., Caral, M., Pochwatko, G.: From demonstration to theory in embodied language comprehension: a review. Cogn. Syst. Res. **29–30**, 66–85 (2014)
37. Iza, M., Konstenius, R.A.: Verb representation and modularity. In: Pérez Miranda, L.A., Izagirre, A. (eds.) Advances in Cognitive Science: Learning, Evolution, and Social Action. University of Basque Country Press (2010)
38. Iza, M., Ezquerro, J.: Computational emotions. In: Liljenström, H. (ed.) Advances in Cognitive Neurodynamics (IV). Springer, Dordrecht (2015)
39. Jackendoff, R.: Possible stages in the evolution of language capacity. Trends. Cogn. Sci. **3**(7) (1999)
40. Jackendoff, R.: Semantics and Cognition. MIT Press (1983)
41. Kahan, T.A., Hely, C.D.: The role of valence and frequency in the emotional Stroop task. Psychon. Bull. Rev. **15**(5), 956–960 (2008)

42. Kanfer, R.: Self-regulatory and other non-ability determinants of skill acquisition. In: Gollwizer, P.M., Bargh, J.A. (eds.) The Psychology of Action: Linking Cognition and Motivation to Behavior, pp. 404–423. Guilford Press, NY (1996)

43. Kaschak, M.P., Borreggine, K.L.: Temporal dynamics of the action-sentence compatibility effect. Q. J. Exp. Psychol. **61**, 883–895 (2008)

44. Kauschke, C., Stenneken, P.: Differences in noun and verb processing in lexical decision cannot be attributed to word form and morphological complexity alone. J. Psycholinguist. Res. **37**(6), 443–452 (2008)

45. Kemmerer, D.: How words capture visual experience: the perspective from cognitive neuroscience. In: Malt, B., Wolff, P. (eds.) Words and the Mind: How Words Capture Human Experience, pp. 289–329. Oxford University Press, Oxford (2010)

46. Kinstch, W.: The Representation of Meaning in Memory. Erlbaum, Hillsdale, NJ (1974)

47. Kohonen, T.: Self-Organizing Maps, 3rd edn. Springer (2001)

48. Lancker Van, Sidtis D.: Has neuroimaging solved the problems of neurolinguistics? Brain Lang. **98**, 276–290 (2006)

49. Langacker, R.W.: Foundations of Cognitive Grammar. Stanford University Press (1987)

50. Larsen, R., Mercer, K.A., Balota, D., Strube, M.J.: Not all negative words slow down lexical decision and naming speed: importance of Word arousal. Emotion **8**(4), 445–452 (2008)

51. Li, P., Shirai, Y.: The acquisition of lexical and grammatical aspect. Mouton de Gruyter, Berlin (2000)

52. Li, P.: Language acquisition in a self-organizing neural network model. In: Quinlan, P.T. (ed.) Connectionist Models of Development, pp. 115–149. Psychology Press (2005)

53. Liljeström, M., Tarkiainen, A., Parviainen, T., Kujala, J., Numminen, J., Hiltunen, J., Laine, M., Salmelin, R.: Perceiving and naming actions and objects. Neuroimage **41**, 1132–1141 (2008)

54. Lugli, L., Baroni, G., Gianelli, C., Borghi, A.M., Nicoletti, R.: Self, others and objects: how this triadic interaction modulates our behavior. Mem. Cogn. **40**, 1373–1386 (2012)

55. MacKay, D.G., Shafto, M., Taylor, J.K., Marian, D.E., Abrams, L., Dyer, J.R.: Relations between emotion, memory and attention: evidence from taboo Stroop, lexical decision and immediate memory tasks. Mem. Cogn. **32**(3), 474–488 (2004)

56. MacWhinney, B.: Emergentist approaches to language. In: Bybee, J., Hopper, P. (eds.) Frequency and Emergency of Linguistic Structure, pp. 449–470. Benjamins, Amsterdam (2001)

57. Mathews, A., MacLeod, C.: Cognitive approaches to emotion and emotional disorders. Annu. Rev. Psychol. **45**, 25–50 (1994)

58. McKoon, G., Ratcliff, R.: Inference during reading. Psychol. Rev. **99**, 440–466 (1992)

59. Mestres-Misse, A., Rodriguez-Fornells, A., Munte, T.F.: Neural differences in the mapping of verb and noun concepts onto novel words. Neuroimage **49**, 2826–2835 (2009)

60. Mukamel, R., Ekstrom, A.D., Kaplan, J.T., Iacoboni, M., Fried, I.: Single-neuron responses during execution and observation of actions. Curr. Biol. **20**(8), 750–756 (2010)

61. Pittermann, J., Pittermann, A., Minker, W.: Handling Emotions in Human-Computer Dialogues. Springer, Berlin (2010)

62. Pons, F., Doudin, P-A., Harris, P.: Enseigner la comprehension es émotions: Quel type de programme pour quel type d'élève?. In Doudin. P.-A., Lafortune, L. (eds.) Enseignement spécialisé: quelle formation pour les enseignants? Paper, Congrès du Réseau d'Échange Francophone. IEEE Signal Processing Society, Geneva (2003)

63. Pulvermüller, F.: Grounding language in the brain. In: de Vega, M., Graesser, A., Glengberg, A.M. (eds.) Symbols, Embodiment and Meaning, pp. 85–116. Oxford University Press, Oxford (2008)

64. Rieman, B.C., McNally, R.J.: Cognitive processing of personally-relevant information. Cogn. Emot. **9**, 325–340 (1995)

65. Rosch, E.H.: On the internal structure of perceptual and semantic categories. In: Moore, T.E., (ed.) Cognitive Development and Acquisition of Language, pp. 111–144. Academic Press, New York (1973)

66. Rosch, E.H.: Linguistic Relativity. In: Silverstein, A. (ed.) Human Communication. Theoretical Explorations. L.E.A, Hillsdale (1974)
67. Rosch, E.H.: Cognitive representations of semantic categories. J. Exp. Psychol. Gen. 104, 192–233 (1975)
68. Rudolph, U., von Hecker: Three principles of explanation: verb schemas, balance and imbalance repair. J. Lang. Soc. Psychol. 25, 377–405 (2006)
69. Rugg, M.D.: The effects of semantic priming and word repetition on event-related potentials. Psychophysiology 22, 642–647 (1985)
70. Taylor, L.J., Zwaan, R.A.: Motor resonance and linguistic focus. Q. J. Exp. Psychol. Sect. A —Hum. Exp. 61, 896–904 (2008)
71. Taylor, L.J., Zwaan, R.A.: Action in cognition: the case of language. Lang. Cogn. 1, 45–58 (2009)
72. Thill, S., Caligiore, D., Borghi, A.M., Ziemke, T., Balsassare, G.: Theories and computational models of affordance and mirror systems: an integrative review. Neurosci. Behav. Rev. 37(3), 491–521 (2013)
73. Thorndyke, P.W.: Conceptual complexity and imagery in comprehension and memory. J. Verbal Learn. Verbal Behav. 14, 359–369 (1975)
74. Tyler, L.K., Randall, B., Stamatakins, E.A.: Cortical differentiation for nouns and verbs depends on grammatical markers. J. Cogn. Neurosci. 20, 1381–1389 (2008)
75. Vanhatalo, U.: The use of questionnaires in exploring synonymy. Lexical knowledge from native speakers to electronic dictionaries. PhD Thesis. Location of document (2005). http://ethesis.helsinki.fi/julkaisut/hum/suoma/vk/vanhatalo
76. Willms, J.L., Shapiro, K.A., Peelen, M.V., Pajtas, P.E., Costa, A., Moo, L.R., Caramazza, A.: Language-invariant verb processing regions in Spanish-English bilinguals. Neuroimage 57, 251–261 (2011)
77. Wray, A., Perkins, M.: The functions of formulaic language: an integrated model. Lang. Commun. 20, 1–28 (2000)
78. Zwaan, R.A., Taylor, L.J.: Seeing, acting, understanding: motor resonance in language comprehension. J. Exp. Psychol.-General 135, 1–11 (2006)