Implications of the Embodied Language: From Learning in Humans to Multisensory Integration in Robots

Michał B. Paradowski[1]

Abstract. A few decades of work in the AI field have focused efforts on developing a new generation of systems which can acquire knowledge via interaction with the world. Yet, until very recently, most such attempts were underpinned by research which predominantly regarded linguistic phenomena as separated from the brain and body. This could lead one into believing that to emulate linguistic behaviour, it suffices to develop 'software' operating on abstract representations that will work on any computational machine. This picture is inaccurate for several reasons, which are elucidated in this paper and extend beyond sensorimotor and semantic resonance. Beginning with a review of research, I list several heterogeneous arguments against disembodied language, in an attempt to draw conclusions for developing embodied multisensory agents which communicate verbally and non-verbally with their environment.

Without taking into account both the architecture of the human brain, and embodiment, it is unrealistic to replicate accurately the processes which take place during language acquisition, comprehension, production, or during non-linguistic actions. While robots are far from isomorphic with humans, they could benefit from strengthened associative connections in the optimization of their processes and their reactivity and sensitivity to environmental stimuli, and in situated human-machine interaction. The concept of multisensory integration should be extended to cover linguistic input and the complementary information combined from temporally coincident sensory impressions.

1 INTRODUCTION

In the 'traditional' view, going back to René Descartes, cognition has been seen as manipulation of symbolic, mental representations, with the brain conceived of as an input-output processor, a problem-solving device running abstract, generalised computational programs which enable us to process incoming data into a perception/interpretation of the outside world. This 'software', separate from the body, was equated with the mind, while the body was regarded as an output system attached to the cognitive processing system, with similar tasks achieved by applying the same underlying motor program to different effectors. The information-processing approach or computer metaphor has become further entrenched over the latter half of the previous century due to the adoption of the digital computer as the platform to run the symbolic computations (Hoffmann *et al.*, n.d.).

Until very recently, most language research has, in a Cartesian manner, traditionally regarded linguistic phenomena as internal, mental, isolationist and amodal (that is, separate and independent from perception, action and emotion systems, and the body); a view endorsed in psychology (e.g. Geschwind 1970; Kintsch 1998), philosophy (e.g. Katz & Fodor 1963; Fodor 1983), and linguistics (e.g. early Chomsky – 1957, 1975; Nowak *et al.* 2002; Jackendoff 2002). For instance, Chomsky's most seminal theories were based on mathematical formalism and saw language as governed by a context-free grammar extended with transformational rules operating on (non-semantic) symbol strings and complemented by morphophonemic rules, with autonomous syntax at the core of the theory of language. The reason why his views for a long time did not go beyond such a perspective should not come as a surprise. His *Syntactic Structures*, which became a revolutionary and foundational work in linguistics, grew out of a series of lecture notes for an audience of undergrad (mainly electrical engineering and maths) students at the MIT. Also, Chomsky's ideas were born at the same time as the establishment of computer science as a distinct academic discipline, the beginnings of computational linguistics, and the founding of AI research, which all shared the dominant idea that thought can be described with formal logic.

The generative school inspired several decades of linguistic thought, and even theories trying to modify or undermine its tenets were still relying on the underlying view of language as a system manipulating abstract symbols. This dualistic view could lead one into believing that in order to credibly emulate linguistic behaviour, it suffices to develop 'software' operating on (i.e. applying combinatorial rules such as Merge and Move to) abstract representations[2] that will work on any computational machine, and that its operations will be implementation-independent, functioning identically regardless of the physical hardware.

2 EMBODIED LANGUAGE IN HUMANS

The dualistic approach just outlined above works to some extent in statistical machine translation, automatic text indexing and retrieval (think e.g. search engines), natural-language interfaces or dialogue systems, but if the system to be developed is to truly mimic human behaviour, the disembodied picture is not very accurate for several reasons. One may be doubtful about modularity and the existence

[1] Inst. of Applied Linguistics, Univ. of Warsaw, ul. Dobra 55, PL-00-312 Warsaw, Poland. E-mail: michal.paradowski@uw.edu.pl.

[2] Understood as terminal symbols, which can—subsequently or concurrently—be equipped with referential, meaning-bearing properties.

of a specifically dedicated innate language acquisition device, but must still take into account the following phenomena and theoretical developments:

1. **lateralization** and **localization** of the language faculty in the brain. Linguistic capabilities have been shown to be limited to certain areas of the cerebrum, as evidenced primarily by various language disorders:
   - receptive aphasia, commonly known as Wernicke's aphasia (Wernicke 1874): damage to the medial temporal lobe de-stroying local language regions and cutting them off from most of the occipital, temporal and parietal regions (cf. e.g. Price 2000; Bookheimer 2002; Damasio *et al.* 2004);
   - expressive aphasia (aka Broca's or agrammatic aphasia; Broca 1861);
   - abnormal language developed in individuals with the left hemisphere removed (Dennis & Whitaker 1976);
   - Specific Language Impairment (SLI), which is unrelated to other developmental disorders, mental retardation, brain in-jury, or deafness (e.g. Joanisse & Seidenberg 1998; Bishop & Snowling 2004; Archibald & Gathercole 2006);
   - other cases of people with normal nonverbal abilities but impaired language, and 'normal' language but cognitive deficits (cf. the classic case studies of individuals with incommensurable linguistic and cognitive capacities: Genie (Curtiss 1981), Laura (Yamada 1990), Clive (Smith 1989), or Christopher (Smith *et al.* 1993).

While these deficits cannot straightforwardly be taken as proof of the *modularity* of language (*cf.* e.g. Calabrette *et al.* 2003; Fodor 2005), they do point to *localisation* of language processes;

2. embodiment of language in neuronal circuitry. FMRI studies have shown '**activation**' of certain brain areas involved in **language processing** (e.g. Osterhout 1997; Hagoort *et al.* 1999; Embick *et al.* 2000; Horwitz *et al.* 2003; Pulvermüller & Assadollahi 2007), with different levels of language processing identified in specific regions, e.g. loci of syntax mainly in left-perisylvian language regions, especially Broca's and Wernicke's areas, but also adjacent neocortical areas, the insula, and subcortical structures including basal ganglia (*cf.* e.g. Ullman 2001; Grodzinsky & Friederici 2006), or phonology in the superior temporal sulcus and anterior superior temporal cortex (*cf.* e.g. Diesch *et al.* 1996; Obleser *et al.* 2006; Uppenkamp *et al.* 2006);
3. genetic influence on language. While mutations of the Foxhead box protein 2 (**FOXP2** gene), deemed to cause a severe speech and language disorder (e.g. Lai *et al.* 2001; Vernes *et al.* 2008; Fisher & Scharff 2009),were initially taken as evidence for a 'language gene', it was later discovered that the protein impacts a wide range of phenotypic features all over the body (including facial motor control) and that the impairments of the family affected with the mutation went beyond language to other cognitive capacities. It is now more believed that it is networks of gene interactions rather than individual genes that have an influence on language (Knopka *et al.* 2009), but the neurobiological influence is there;
4. many Universal Grammar-based constraints now being reinterpreted as **learning** and **processing constraints**. That is, the difficulty in the acquisition of certain aspects of language are being accounted for by their complexity, the computational load under which the user/learner operates, his/her memory and attention limitations, or ease of access to representations (*cf.* e.g. Wakabayashi 2002; Van Hell & De Groot 1998; Wątorek 2008);
5. maturation and the **critical/sensitive period** (but consider e.g. Marinova-Todd *et al.* 2000 for a contradictory view);
6. the Chomskyan **competence** vs. **performance** distinction (Chomsky 1965), explaining mistakes in (originally native) language users' output (i.e., their actual deployment of the linguistic capacity) attributable to such psychosomatic states and factors affecting them as fatigue, tedium, intoxication, drugs, sudden changes of mind, haste, inattention, or external distractions;
7. interaction between (context-bound) language comprehension and production, and sensorimotor activation, manifested in both directions by:
   - **motor resonance** observed in linguistic (Lakoff & Johnson 1980; Lakoff 1987), behavioural (primarily with priming modulating motor performance; e.g. Tanenhaus *et al.* 1995; Gentilucci *et al.* 2000; Spivey *et al.* 2001; Glenberg & Kaschak 2002; Glover *et al.* 2004; Buccino *et al.* 2005; Boulenger *et al.* 2008; Nazir *et al.* 2008; Frak *et al.* 2010; for grammar *cf.* Madden & Zwaan 2003; Bergen & Wheeler 2010), neuroimaging and TMS studies[3] (e.g. Zatorre *et al.* 1992; Fadiga *et al.* 2002; Tettamanti *et al.* 2008; Fischer & Zwaan 2008; Kemmerer *et al.* 2008; Boulenger *et al.* 2009; Willems *et al.* 2010; for activation in visual areas *cf.* Martin *et al.* 1996; Pulvermüller & Hauk 2006; Simmons *et al.* 2007; in the olfactory cortex *cf*. González *et al.* 2006);
   - **semantic resonance** (brain language areas getting activated during sensorimotor action; Bonda *et al.* 1994; Pulvermüller *et al.* 2005; Rueschemeyer *et al.* 2010);

---

[3] Somewhat importantly, motor resonance was not observed when the stimuli were used in idiomatic contexts (Rueschemeyer *et al.* 2010a) or metaphorical ones. Regarding the latter, Raposo *et al.* (2009) found activity in the pre- and motor cortex for literal-only usages of arm- and leg-related Vs, while Bergen *et al.* (2007) likewise demonstrated that visual imagery is triggered in sentence comprehension tasks (where general words of motion were employed) only where the utterances have literal spatial meaning. However, the picture is not completely clear-cut. This year, Lacey *et al.* (2012) showed that textural metaphors do activate parietal operculum regions important to the sense of touch. To explain this discrepancy, one could posit a qualitative difference between 'directly' embodied sensory experiences (e.g. texture or temperature) and more 'indirect' ones such as those grounded in visual perception. The former are more 'primary':

i) sensed earliest – already in the womb, tactition being the first sense that begins to develop before 8 weeks gestational age together with the emergence of the nervous system (Montagu 1978), before taste and smell (14 weeks g.a.), hearing (16 weeks g.a.; Shahidullah & Hepper 1992) or vision (week 18 onwards),

ii) available in more 'primitive' organisms without vision or hearing,

iii) perceptible during half-sleep, and

iv) impacting our bodily functioning more strongly (the somatic reaction to extremely high or low temperatures, pressure or skin irritation is more likely to be stronger than e.g. to an unpleasant sight or sound).

This might account for the lack of activation in visual cortical areas.

- verbalization of memory facilitated when assuming the original body position during recall (Dijkstra *et al.* 2007), linguistic tasks expedited when accompanied by action (Rieser *et al.* 1994), and sensorimotor experiences intertwined with cognition in episodic memory (Pfeifer 2011);
- faster comprehension of depictions of spatial associations than of descriptions of spatial *dis*sociations (Glenberg *et al.* 1987); speedier recognition of words with 'body-object interaction' than of ones without (Siakaluk *et al.* 2008);
- semantic interference and facilitation in the Stroop effect (longer RTs needed to name colour names written in incongruent ink hue; Jaensch 1929; Stroop 1935);
- clinical studies indicating that processing of action concepts degrades if action- or vision-related brain areas are lesioned in motor neuron diseases (Damasio *et al.* 1996; Bak *et al.* 2001; Neininger & Pulvermüller 2003) and semantic dementia (Pulvermüller *et al.* 2010);
- comprehension of action words deteriorating after loss of procedural knowledge (*cf.* Boulenger *et al.* 2008 on Parkinson's disease patients; also Bak *et al.* 2006);

8. parallel emergence of speech and gesture in infancy (Iverson & Thelen 1999);
9. co-speech gesture reducing cognitive load (Goldin-Meadow *et al.* 2001), and indications of a dual-task advantage for bimodal (signed-spoken) language production (i.e., production of code-blends, with elements of the signed and spoken languages appearing simultaneously; Kaufmann & Kaul 2012); or
10. Conceptual Blending theory (Fauconnier & Turner 2002) explaining language creativity as a semantic process operating on the output of perception and interaction with the world to create new structures.

Thus, independently of theoretical persuasion, without taking into account both the architecture of the human brain, and embodiment—the interaction of the language faculty with the sensory apparatus and motor system—it is unrealistic to replicate accurately the processes which take place during language acquisition, comprehension, or production, or during non-linguistic actions. Cognitive mechanisms are synergistically intertwined with affective and somatic components, and largely inseparable (Ziemke 2011).

3 THE COROLLARIES FOR ROBOTICS

Since the official launch of AI as a new research discipline at the seminal Dartmouth conference in 1956, much of work in the field has been driven by the 'Physical Symbol Hypothesis' (Newell & Simon 1976): trying to construct systems that would possess or build internal, symbolic representations of objects and relations in the outside *world*—in other words, a "world model"—which usually had little to do with their hardware, sensorimotor *experience*, or current context, but were instead characterised by precisely defined states and finite lists of acceptable commands (Wang 2009:2*f.*). Under such a functionalist approach, the body is merely a platform on which cognitive operations are running. In some areas, such closed systems were able to achieve spectacular feats, for instance in defeating world chess champions.

Chess, however, is a formal game, set in a virtual world with discrete states, positions, and licit moves, a game involving complete information, and a static one: no move means no change, and the inventory of legitimate operations remains constant (Pfeifer & Scheier 1999:58*ff.*). This is quite unlike what usually happens in the real world. Hence, the last two and a half decades have witnessed recurrent appeals for situated, embodied autonomous systems actively and directly interacting with the world around (*cf. op. cit.*; Brooks 1991; Varela *et al.* 1991) and constructing knowledge via this dynamic enactment (the active learning being qualitatively different from statistical machine learning; *cf.* e.g Froese 2009; Vernon 2010). Evidently robots, even anthropomorphic ones, are far from isomorphic with humans in terms of both the 'brain' and the rest of the body, including the input and output devices (sensors and actuators). Also, as one reviewer rightly remarks, in the language technology field priority is not necessarily to make a machine as humanlike as possible, with the same architecture; rather, it is to make the machine so that it does things on a level comparable to humans (or, I would add, surpassing that) – in other words, to achieve similar—or better—functionality in terms of mode, scope, or scale. Or, going completely beyond the anthropocentric GOFAI perspective (Haugeland 1985; *cf.* Wang 2008), since passing the Turing Test is not a *sine qua non* of being intelligent, as acknowledged by the test's designer himself (Turing 1950). This, however, means that robust artificial cognitive agents can bypass the human limitations inherent in most of the above points (just as they could overcome some contingencies resulting from the material properties of the human brain and bodily features such as synaptic speed and efficiency, the physical characteristics of the vocal tract, the auditory perception system, or muscular flexibility). Nevertheless, they could still benefit from strengthened associative connections owing to the motor and semantic resonance in both the optimization of their processes, and reactivity and sensitivity to environmental stimuli, across a range of tasks:

(i) in grounded language understanding (*cf.* e.g. Glenberg & Kaschak 2002; Feldman & Narayanan 2004; Gallese & Lakoff 2005; Sato *et al.* 2008), where structuring the environment acts as scaffolding and all inputs contribute to evidential support,

(ii) in automated articulation-based speech recognition (utilising motor information, i.e. combining spoken input with visual data—e.g. the shape of the speakers lips—and maybe even data such as strength of the incoming airstream),

(iii) while learning about context-dependent phenomena in the surrounding world (e.g. action sequences and argument structure in construction grammar; *cf.* Dominey 2007; since embodiment plays a constitutive role in the process of cognition; Vernon 2010), or in the process of language acquisition in general (because language—at least in the initial stages—is acquired by situated embodied direct engagement with the world, and not just passive perception, e.g. watching television; *cf.* e.g. Steels 2009),

(iv) to help with storage and retrieval due to the benefits of episodic memory,

(v) to support action prediction, planning and anticipation (Koelewijn *et al.* 2008; Stapel *et al.* 2010; van Elk *et al.* 2010), including prediction of the next sensory feedback,

(vi) to support action execution (with linguistic input making the actor better aware of the affordances, i.e. physically feasible action possibilities), and

(vii) to reinforce feedback in ‘soft robotics’ and morphological computation, where there is no clear separation between the controller (or orchestrator) and the hardware (morphology), and the tasks are distributed between the brain, body, and environment (*cf.* e.g. Paul 2004; Pfeifer 2011; which also has the aim of off-loading computation; Di Paolo 2009);

(viii) in cognitive developmental robotics, aiming at understanding human cognitive developmental processes by synthetic or constructive approaches (Asada *et al.* 2009, Asada 2011, Ishiguro *et al.* 2011);

(ix) in common grounding and alignment, which are crucial for fruitful situated human-machine interaction, and which are another area where sensory experience must be coordinated with linguistic interaction.

Principally, if our goal were to create machines which do things on a *comparable level* to—or surpassing—humans, we could do away with attempts at embodying them in human-inspired ways (Taivo Lints, p.c., 31 May 2012) – they could function perfectly well with totally nonhuman kinds of embodiment (different ‘bodies’, different sensors and effectors, different internal architectures... or even with embodiment in a virtual world; Bringsjord *et al.* 2008; Goertzel *et al.* 2008). Given the role played by the morphology of the sensory apparatus and the architecture of the sensorimotor loop in shaping and structuring the information that reaches the controller, and thereby in concept formation, it would anyway be difficult for a machine to form the same concepts, categories and behaviours as us without having comparable morphology (as remarked e.g. by Barsalou 1999 or Lakoff & Johnson 1998). However, if our goal is to have machines ‘thinking’ and behaving in a way *compatible* with ours—which is a highly practical and desirable goal—then it is of high importance for them to develop, learn and function in a similar “experience space” (Taivo Lints, p.c.; *cf.* also Wang 2009:5).

The requirement that the behaviour, perception and conceptual apparatus of artificial intelligent agents be grounded in their experience of their *own* interaction with the outside world at once means that their concepts and categories need not necessarily rely on the same minimal constituents and grammatical categories as have been externally identified and defined in linguistics. Instead, the gradually emergent categories are more likely to be intrinsically meaningful behaviours and affordances (see also Kuniyoshi *et al.* 2004), action-oriented rather than orbocentric (Hoffmann & Pfeifer 2011). For instance, to a robot who has never kicked or observed anyone kick anything but footballs, the minimal unit of meaning may be <kick a ball> rather than <kick> alone (although this does not rule out the possibility of extrapolation and abstraction should a relevant opportunity arise).[4] Similarly, irrespective of whether the input is expressed using [$_{NP}$ kicking a ball] or [$_{VP}$ kick a ball], it should activate the same action schema.

4 TOWARDS A BROADER DEFINITION OF MULTISENSORY INTEGRATION

In order to form a meaningful experience and construct coherent, reliable and robust representations of the surrounding world, the human brain combines prior knowledge with sensory input arriving from various modalities and integrates these at multiple levels of the neuraxis. This serves to maximize the efficiency of everyday performance and learning, enhancing the salience of the events, helping increase the detection and identification of the external stimuli, disambiguate them, compensate for incomplete information, and shorten reaction times. In view of the inseparability of language and the body, the concept of multisensory integration—whether in natural or artificial cognitive agents—should be extended and cover both the linguistic input and the complementary information that the brain combines from temporally coincident sensory impressions. This does not mean that we should ‘dumb down’ the statistical processes where they operate successfully; instead, where the input stream in one channel is too noisy, turning on auxiliary channels and interacting with the environment in an active manner may generate ancillary data and help e.g. disambiguate the signal and take the right decision (see also Pfeifer & Scheier 1997; Beer 2003).[5] An added benefit would then be significantly reduced programming costs.

ACKNOWLEDGMENTS

The author wishes to thank Noam Chomsky, Anna Esposito, Richard Littauer, Gary Lupyan, Vincent C. Müller, Michael Pleyer, and Luc Steels for invaluable commentary, discussion and bibliographical references. Naturally, willingness to comment does not imply endorsement; all the usual disclaimers apply. An earlier version of this paper was presented at the Symposium “Linguistic and Cognitive Approaches to Dialogue Agents”, AISB/IACAP World Congress 2012 in honour of Alan Turing, University of Birmingham, 2 July 2012.

---

[4] See for instance the POETICON++ project (Robots need Language: A computational mechanism for generalisation & generation of new behaviours in robots; `http://www.poeticon.eu/`).

[5] One consequence for humans may be that the role of kinaesthetic modality, traditionally largely believed to dominate in children, but be negligible in adults (*cf.* e.g. Barbe & Milone 1981; Felder & Spurlin 2005), should be reassessed, as the effectiveness may be demonstrated of ‘learning-by-doing’ and task-based approaches to language learning and teaching where the students have to use their bodies (e.g. when acquiring novel lexis via common cookery classes). This at once is an implication for language learning. Another one is that given the role mirror neurons play in our interactions with and understanding of the world, the coupling between language and actions implies a positive asset for computer games in fostering language development.

REFERENCES

[1] Archibald, L.M.D., Gathercole, S.E.: Prevalence of SLI in Language Resource Units. *J Res Spec Educ Needs* 6(1), 3–10 (2006).

[2] Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., Ogino, M., Yoshida, C.: Cognitive developmental robotics: a survey. *IEEE Transactions on Autonomous Mental Development* 1(1), 12--34 (2009).

[3] Asada, M.: Can cognitive developmental robotics cause a paradigm shift? In: Krichmar, J.L., Wagatsuma, H. (Eds) *Neuromorphic and Brain-Based Robots: Trends and Perspectives*, pp. 251–273, Cambridge: Cambridge University Press (2011).

[4] Bak, T.H., O'Donovan, D.G., Xuereb, J.H., Boniface, S., Hodges, J.R.: Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain* 124, 103–120 (2001).

[5] Bak, T.H., Yancopoulou, D., Nestor, P.J., Xuereb, J.H., Spillantini, M.G., Pulvermuller, F., Hodges, J.R.: Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain* 129(2), 321–332 (2006).

[6] Barbe, W.B., Milone, M.N.: What we know about modality strengths, *Educational Leadership* 38(5), 378–80 (1981).

[7] 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).

[8] Bergen, B., Lindsay, S., Matlock, T., Narayanan, S.: Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cogn Sci* 31, 733–764 (2007).

[9] Bergen, B., Wheeler, K.B.: Grammatical aspect and mental simulation. *Brain Lang* 112, 150–158 (2010).

[10] Bishop, D.V.M., Snowling, M.J.: Developmental dyslexia and specific language impairment: same or different? *Psychol Bull* 130, 858–88 (2004).

[11] Bonda, E., Petrides, M., Frey, S., Evans, A.: Frontal cortex involvement in organized sequences of h and movements: Evidence from a positron emission topography study. *Soc Neurosci Abstr* 20, 353 (1994).

[12] Bookheimer, S.: Functional MRI of l anguage: New approaches to understanding the cortical organization of semantic processing. *Ann Rev Neurosci* 25, 151–188 (2002).

[13] Boulenger, V., Hauk, O., Pulvermüller, F.: Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cereb Cortex* 19(8), 1905–1914 (2009).

[14] Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., Nazir, T.A.: Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia* 46(2), 743–756 (2008).

[15] Broca, P. : Remarques sur le siège de la faculté de la parole articulée, suivies d'une observation d'aphémie (perte de parole). *Bull Soc Anatom (Paris)* 36, 330–357 (1861).

[16] Calabretta, R., Di Ferdinando, A., Wagner, G. P., Parisi, D.: What does it take to evolve behaviorally complex organisms? *BioSystems* 69, 245–262 (2003).

[17] Chomsky, N. *Syntactic structures*. The Hague: Mouton (1957).

[18] Chomsky, N. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press (1965).

[19] Chomsky, N. *Reflections on language*. New York: Pantheon (1975).

[20] Chomsky, N. *Rules and Representations*. New York: Columbia University Press (1980).

[21] Chomsky, N. Three factors in language design. *Ling Inq* 36(1), 1-22 (2005).

[22] Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. A neural basis for lexical retrieval. *Nature*, 380, 499–505 (1996).

[23] Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. Neural systems behind word and concept retrieval. *Cognition* 92:179–229 (2004).

[24] Dennis, M., Whitaker, H. Language acquisition following hemidecortication: Linguistic superiority of the left over the right hemisphere. *Brain Lang* 3, 404-433 (1976).

[25] Di Paolo, E. From sensorimotor coordination to enaction: Agency, sense-making and sociality as horizons for e mbodied cognition. 1st EUCogII Members' Conference "Challenges for artificial cognitive systems", University Medical Center Hamburg-Eppendorf (2009, Oct 10).

[26] Di Sciullo, M., Boeckx, C. *The Biolinguistic Enterprise. New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press (2011).

[27] Diesch, E., Eulitz, C., Hampson, S., & Ross, B. The neurotopography of vow els as mirrored by evoked magnetic field measurements. *Brain Lang* 53(2), 143–168 (1996).

[28] Dijkstra, K., Kaschak, M., Zwaan, R.: Body posture facilitates retrieval of autobiographical memories. *Cognition* 102, 139–49 (2007).

[29] Dominey, P.F.: Spoken language and vision for adaptive human-robot cooperation. In: Hackel, M. (ed.), *Humanoid robotics*. ARS International, Vienna (2007).

[30] Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., Sakai, K.L.: A syntactic specialization for Broca's area. *Proc Natl Acad Sci USA* 97, 6150–6154 (2000).

[31] Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G.: Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *Eur J Neurosci* 15(2), 399–402 (2002).

[32] Fauconnier, G., Turner, M.: *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books (2002).

[33] Felder, R.M., Spurlin, J.: Applications, reliability and validity of the Index of Learning Styles. *Int J Engng Ed* 21(1), 103–12 (2005).

[34] Feldman, J., Narayanan, S.: Embodied meaning in a neural theory of language. *Brain Lang* 89(2), 385–392 (2004).
[35] Fischer, M.H., Zwaan, R.A.: Embodied language: A review of the role of motor system in language comprehension. *Q J Exp Psychol* 61, 825–850 (2008).
[36] Fisher, S.E., Scharff, C.: FOXP2 as a molecular window into speech and language. *Trends Genet* 25(4), 166–177 (2009).
[37] Fodor, J.A.: *The Modularity of Mind: An essay on faculty psychology*. MIT Press, Cambridge, MA (1983).
[38] Fodor, J.A.: Reply to Steven Pinker "So How Does the Mind Work?" *Mind & Language* 20, 25–32 (2005).
[39] Frak, V., Nazir, T., Goyette, M., Cohen, H., Jeannerod, M.: Grip force is part of the semantic representation of manual action verbs. *PLoS ONE* 5(3), e9728. doi:10.1371/journal.pone.0009728 (2010).
[40] Gallese, V., Lakoff, G.: The brain's concepts: The role of the sensory–motor system in reason and language. *Cogn Neuropsychol* 22, 455–479 (2005).
[41] Gentilucci, M., Benuzzi, F., Bertonali, L., Daprati, E., Gangitano, M.: Language and motor control. *Exp Brain Res* 133(4), 468–490 (2000).
[42] Geschwind, N.: The organization of language and the brain. *Science* 170(961), 140–144 (1970).
[43] Glenberg, A.M., Kaschak, M.P.: Grounding language in action. *Psychon Bull Rev* 9(3), 558–565 (2002).
[44] Glenberg, A., Meyer, M., Lindem, K.: Mental models contribute to foregrounding during text comprehension. *J Mem Learning* 26: 69–83 (1987).
[45] Glover, S., Rosenbaum, D.A., Graham, J., Dixon, P.: Grasping the meaning of words. *Exp Brain Res* 154 (1), 103–108 (2004).
[46] González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., Ávila, C.: Reading *cinnamon* activates olfactory brain regions. *Neuroimage* 32(2), 906–912 (2006).
[47] Grodzinsky, Y., Friederici, A.D.: Neuroimaging of syntax and syntactic processing. *Curr Opin Neurobiol* 16(2), 240–246 (2006).
[48] Hagoort, P., Ramsey, N.F., Rutten, G.J.M., van Rijen, P.C.: The role of the left anterior temporal cortex in language processing. *Brain Lang* 69, 322–325 (1999).
[49] Hauser, M.D., Chomsky, N., Fitch, W.T.: The language faculty: What is it, who has it, and how did it evolve? *Science* 298(5598), 1569–1579 (2002).
[50] Horwitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, J., Zilles, K., Braun, A.R.: Activation of Broca's area during the production of spoken and signed language: A combined cytoarchitectonic mapping and PET analysis. *Neuropsychologia* 41, 1868–1876 (2003).
[51] Ishiguro, H., Minato, T., Yoshikawa, Y., Asada, M.: Humanoid Platforms for Cognitive Developmental Robotics. *Intl J Humanoid Robotics* 8(3), 391–418 (2011).
[52] Iverson, J., Thelen, E.: Hand, mouth, and brain: The dynamic emergence of speech and gesture. *J Consc Stud* 6(11-12), 19–40 (1999).
[53] Jackendoff, R.: *Foundations of Language*. Oxford University Press, Oxford (2002).
[54] Jaensch, E.R : *Grundformen menschlichen Seins*. Berlin: Otto Elsner (1929).
[55] Joanisse, M.F., Seidenberg, M.S.: Specific language impairment: A deficit in grammar or processing? *Trends Cogn Sci* 2, 240–247 (1998).
[56] Katz, J. J., Fodor, J.: The structure of a semantic theory. *Lang* 39, 170–210 (1963).
[57] Kemmerer, D., Castillo, J.G., Talavage, T., Patterson, S., Wiley, C.: Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang* 107, 16–43 (2008).
[58] Kemmerer, D., Gonzalez-Castillo, J.: The two-level theory of verb meaning: an approach to integrating the semantics of action with the mirror neuron system. *Brain Lang* 112, 54–76 (2010).
[59] Kintsch, W.: *Comprehension: A Paradigm for Cognition*. Cambridge University Press, Cambridge, MA (1998).
[60] Knopka G., Bomar, J.M., Winden, K., Coppola, G., Jonsson, Z.O., Gao, F., Peng, S., Preuss, T.M., Wohlschlegel, J.A., Geschwind, D.H.: Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* 462, 213–217 (2009).
[61] Koelewijn, T., van Schie, H.T., Bekkering, H., Oostenveld, R., Jensen, O.: Motor-cortical beta oscillations are modulated by correctness of observed action. *Neuroimage* 40, 767–775 (2008).
[62] Lacey, S., Stilla, R., Sathian, K.: Metaphorically feeling: Comprehending textural metaphors activates somatosensory cortex. *Brain Lang* 120(3), 416–421 (2012).
[63] Lai, C.S.L., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., Monaco, A.P.: A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413(6855), 519–523 (2001).
[64] Lakoff, G.: *Women, Fire, and Dangerous Things. What categories reveal about the mind*. University of Chicago Press, Chicago (1987).
[65] Lakoff, G., Johnson, M.: *Metaphors we Live By*. University of Chicago Press, Chicago (1980).
[66] Lamendella, J.T.: General principles of neurofunctional organization and their manifestations in primary and non-primary language acquisition. *Language Learning* 27: 155–96 (1977).
[67] Lenneberg, E.H.: *Biological Foundations of Language*. Wiley, New York (1967).
[68] Madden, C.J., Zwaan, R.A.: How does verb aspect constrain event representations? *Mem Cognit* 31, 663–672 (2003).
[69] Marinova-Todd, S., Marshall, D. Snow, C.: Three misconceptions about age and L2 learning. *TESOL Quar* 34, 9–34 (2000).
[70] Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V.: Neural correlates of category-specific knowledge. *Nature* 379, 649–652 (1996).

[71] Meader, C.L., Muyskens, J.H.: *Handbook of Biolinguistics*. Wiley, New York (1950).
[72] Montagu, A.: *Touching: The Human Significance of the Skin*. New York: Harper & Row (1978).
[73] Nazir, T.A., Boulenger, V., Roy, A.C., Silber, B.Y., Jeannerod, M., Paulignan, Y.: Language-induced motor perturbations during the execution of a reaching movement. *Q J Exp Psychol* 61(6), 933–943 (2008).
[74] Neininger, B., Pulvermüller, F.: Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia* 41(1), 53–70 (2003).
[75] Nowak, M.A., Komorova, N. L., Niyogi, P.: Computational and evolutionary aspects of language. *Nature* 417, 611–617 (2002).
[76] Obleser, J., Boecker, H., Drzezga, A., Haslinger, B., Hennenlotter, A., Roettinger, M., Eulitz, C., Rauschecker, J.P.: Vowel sound extraction in anterior superior temporal cortex. *Hum Brain Mapp* 27(7), 562–571 (2006).
[77] Osterhout, L.: On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain Lang* 59, 494–522 (1997).
[78] Paul, C.: Morphology and computation. *Proc Int Conf Simulation Adaptive Behavior: From animals to animats*, pp. 33–38. MIT Press, Cambridge, MA (2004).
[79] Penfield, W., Roberts, L.: *Speech and Brain Mechanisms*. Atheneum Press, New York (1959).
[80] Pfeiffer, R.: The emergence of cognition from the interaction of brain, body, and environment. 4th EUCogII Members' Conference "Embodiment – Fad or Future?", Anatolia College, Thessaloniki (2011, Apr 11).
[81] Piatelli-Palmarini, M.: Evolution, selection and cognition: from "learning" to parameter setting in biology and in the study of language. *Cognition* 31, 1–44 (1989).
[82] Postle, N., McMahon, K.L., Ashton, R., Meredith, M., de Zubicaray, G.I.: Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *Neuroimage* 43, 634–644 (2008).
[83] Price, C.J.: The anatomy of language: contributions from functional neuroimaging. *J Anat* 197, 335–359 (2000).
[84] Pulvermüller, F., Assadollahi, R.: Grammar or serial order? Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. *J Cogn Neurosci* 19(6), 971–980 (2007).
[85] Pulvermüller, F., Hauk, O.: Category-specific processing of color and form words in left fronto-temporal cortex. *Cereb Cortex* 16(8), 1193–1201 (2006).
[86] Pulvermüller, F., Hauk, O., Nikulin, V.V., Ilmoniemi, R.J.: Functional links between motor and language systems. *Eur J Neurosci* 21(3), 793–797 (2005).
[87] Pulvermüller, F., Pye, E., Dine, C., Hauk, O., Nestor, P., Patterson, K.: The word processing deficit in semantic dementia: All categories are equal but some categories are more equal than others. *J Cogn Neurosci* 22(9), 2027–2041 (2010).
[88] Raposo, A., Moss, H.E., Stamatakis, E.A., Tyler, L.K.: Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47, 388–396 (2009).
[89] Rieser, J., Garing, A., Young, M.: Imagery, action and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Dev* 45, 1043–1056 (1994).
[90] Rueschemeyer, S.-A., Lindemann, O., van Rooij, D., van Dam, W., Bekkering, H.: Effects of intentional motor actions on embodied language processing. *Exp Psychol* 57(4), 260–266 (2010).
[91] Shahidullah, S., Hepper, P.G. Hearing in the Fetus: Prenatal Detection of Deafness. *Intl J Prenatal and Perinatal Studies* 4(3/4), 235–240 (1992).
[92] Siakaluk, P., Pexman, P.. Aguilera, L., Owen, W., Sears, C.: Evidence for the activation of sensorimotor information during visual word recognition: The body-object interaction effect. *Cognition* 106, 433–443 (2008).
[93] Sato, M., Mengarelli, M., Riggio, L., Gallese, V., Buccino, G.: Task related modulation of the motor system during language processing. *Brain Lang* 105(2), 83–90 (2008).
[94] Simmons, W.K., Ramjee, V., Beauchamp, M.S., McRae, K., Martin, A., Barsalou, L.W.: A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45(12), 2802–2810 (2007).
[95] Smith, N.V., Tsimpli, I.-M., Ouhalla, J.: Learning the impossible: The acquisition of possible and impossible languages by a polyglot savant. *Lingua* 91, 279–347 (1993).
[96] Spivey, M.J., Tyler, M.J., Eberhard, K.M., Tanenhaus, M.K.: Linguistically mediated visual search. *Psychol Sci* 12, 282–286 (2001).
[97] Stapel, J.C., Hunnius, S., van Elk, M., Bekkering, H.: Motor activation during observation of unusual vs. ordinary actions in infancy. *Soc Neurosci* 5, 451–460 (2010).
[98] Steels, L.: The origins and evolution of languages: Darwin's unsolved mystery. International workshop "150 Years after Darwin: From Molecular Evolution to Language", Inst for Cross-Disciplinary Physics and Complex Systems, Palma de Mallorca (2009, Nov 26).
[99] Stein, B.E., London, N., Wilkinson, L.K., Price, D.D.: Enhancement of Perceived Visual Intensity by Auditory Stimuli: A Psychophysical Analysis. *J Cogn Neurosci* 8(6), 497–506 (1996).
[100] Stewart, J., Gapenne, O., Di Paolo, E.A. (Eds) *Enaction: Toward a New Paradigm for Cognitive Science.* Cambridge, MA: MIT Press (2010).
[101] Stroop, J.R.: Studies of interference in serial verbal reactions. *J Exp Psychol* 18 (6), 643–662 (1935).
[102] Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K., Sedivy, J.C.: Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 632–634 (1995).

[103] Tettamanti M., Manenti, R., Della Rosa, P.A., Falini, A., Perani, D., Cappa, S.F., Moro, A.: Negation in the brain: Modulating action representations. *Neuroimage* 43, 358–367 (2008).

[104] Ullman, M.T.: A neurocognitive perspective on language: The declarative/procedural model. *Nature Rev Neurosci* 2(10), 717–726 (2001).

[105] Uppenkamp, S., Johnsrude, I.S., Norris, D., Marslen-Wilson, W., Patterson, R.D.: Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* 31(3), 1284–1296 (2006).

[106] van Elk, M., van Schie, H.T., Zwaan, R.A., Bekkering, H.: The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *Neuroimage* 50, 665–677 (2010).

[107] Vernes, S.C., Nicod, J., Elahi, F.M., Coventry, J.A., Kenny, N., Coupe, A.M., Bird, L.E., Davies, K.E., Fisher, S.E.: Functional genetic analysis of mutations implicated in a human speech and language disorder. *Hum Mol Genet* 15(21), 3154–3167 (2006).

[108] Vernon, D.: Cognitive development and the iCub humanoid robot. 2nd EUCogII Members' Conference "Development of Cognition in Artificial Agents", Univ Zürich (2010, Jan 29).

[109] Wernicke, C.: *Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis*. Kohn & Weigert, Breslau (1874).

[110] Willems, R.M., Hagoort, P., Casasanto, D.: Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychol Sci* 21, 67–74 (2010).

[111] Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A.: Lateralization of phonetic and pitch discremination in speech processing. *Science* 256, 846–849 (1992).

[112] Ziemke, T.: Human embodied cognition. Scientific evidence and technological implications. 4th EUCogII Members' Conference "Embodiment – Fad or Future?", Thessaloniki: Anatolia College (2011, Apr. 12).

[113] Zwaan, R.A., Taylor, L J., de Boer, M.: Motor resonance as a function of narrative time: further tests of the linguistic-focus hypothesis. *Brain Lang* 112, 143–149 (2010).