Affective Computing and Interaction:

Psychological, Cognitive and Neuroscientific Perspectives

Didem Gökçay
Middle East Technical University, Turkey

Gülsen Yildirim
Middle East Technical University, Turkey

Information Science REFERENCE

INFORMATION SCIENCE REFERENCE
Hershey · New York

Director of Editorial Content: Kristin Klinger
Director of Book Publications: Julia Mosemann
Acquisitions Editor: Lindsay Johnston
Development Editor: Dave DeRicco
Publishing Assistant: Milan Vracarich Jr.
Typesetter: Milan Vracarich Jr., Casey Conapitski
Production Editor: Jamie Snavely
Cover Design: Lisa Tosheff

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Affective computing and interaction : psychological, cognitive, and neuroscientific perspectives / Didem Gokcay and Gulsen Yildirim, editors.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-61692-892-6 (hardcover) -- ISBN 978-1-61692-894-0 (ebook) 1. Human-computer interaction. 2. Human-machine systems. 3. Affect (Psychology)--Computer simulation. I. Gokcay, Didem, 1966- II. Yildirim, Gulsen, 1978-
QA76.9.H85A485 2011
004.01'9--dc22
2010041639

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 10
The Role of Affect and Emotion in Language Development

Annette Hohenberger
Middle East Technical University, Turkey

ABSTRACT

In this chapter, language development is discussed within a social-emotional framework. Children's language processing is gated by social and emotional aspects of the interaction, such as affective prosodic and facial expression, contingent reactions, and joint attention. Infants and children attend to both cognitive and affective aspects in language perception ("language" vs. "paralanguage") and in language production ("effort" vs. "engagement"). Deaf children acquiring a sign language go through the same developmental milestones in this respect. Modality-independently, a tripartite developmental sequence emerges: (i) an undifferentiated affect-dominated system governs the child's behavior, (ii) a cognitive and language-dominated system emerges that attenuates the affective system, (iii) emotional expression is re-integrated into cognition and language. This tightly integrated cognitive-affective language system is characteristic of adults. Evolutionary scenarios are discussed that might underlie its ontogeny. The emotional context of learning might influence the course and outcome of L2-learning, too.

INTRODUCTION

Language is a central aspect of human cognition. In cognitive science with its predominant computational perspective, cognition and language, as its core module, have been viewed as research areas that are strictly separated from emotion (Davidson, 2000; Smith & Kosslyn, 2007, chapter 8; Harris, Berko Gleason, & Aycicegi, 2006; Ochsner & Barrett, 2001; Niedenthal, 2007). Affect and emotion, which are inherently subjective processes and feelings, did not fit into the concept of the human mind that was thought to be governed by universal abstract symbols and rules. The two spheres were therefore considered orthogonal to each other, if not oppositional. Their interplay was

DOI: 10.4018/978-1-61692-892-6.ch010

only poorly understood (Forgas, 2008). Emotion was even defamed as a potentially destructive and subversive power that undermined the functioning of the rational mind (ibid.). However, following the 'cognitive revolution' in the middle of the 20th century, we now witness an 'emotion revolution' (Harris et al., 2006, p. 2258; Caldwell-Harris, 2008) in contemporary times. Emotion is no longer the pariah of cognitive science but is now becoming an increasingly respected partner of cognition.[1]

In this chapter, the role of affect and emotion in language development is surveyed. The intimate link between affective and language development has nowhere been more dramatically established than in the crude historical "experiments" in search for the "proto-language" of mankind. It is historically bequeathed that several emperors, namely the Pharaoh Psammetic, the Staufer King Frederic II, and the Scottish King Jacob IV arranged for rigorous experiments on newborns which they had deprived of any language and human companionship in order to find out what language they would develop. This language should then be considered the human proto-language. These experiments all failed: the poor infants either died (as in Frederic II's case) or uttered only some sparse proto-words (which, however, led Psammetic to conclude that Phrygian must be the proto-language and Jacob that it was Hebrew). Thus, the most basic condition which must be met for any language-learning to be possible at all is human companionship and the willingness to communicate, i.e, "human and humane contact." (Goldin-Meadow, 2003, p. 48; cf. also Hohenberger, 2004) A similar claim has been made for feral children grown up without any social and affective interaction (Kuhl, 2007, p. 116).

Although the link between affect and language is now well established, we still do not know the exact nature of this link. In this chapter, I will present various studies and views on the link between affect and language in infant and child development. In order to provide a neuroscience backdrop to the present topic, neuropsychological studies are presented to the reader in the appendix.

THE SOCIAL-EMOTIONAL FRAMEWORK OF LANGUAGE DEVELOPMENT

The study of the social origins of language and cognitive development has its roots in the work of Bruner (1975, 1983) and Vygotsky (1962). They argue that all learning, i.e., also language learning, is rooted in social interaction, most notably between parents and their children (cf. also Tamis-LeMonda & Rodriguez, 2008; Paavola, 2006). Recently, the role of social interaction in cognitive and language development has received again much attention (Tomasello, 2003, 2008), in particular its neural underpinnings (Striano & Reid, 2006; Johnson, 2007, 2008; Grossmann & Johnson, 2007; Kuhl, 2007). This shift in focus from a predominantly computationally to a more socially oriented view of brain processes is relevant for our investigation here insofar as the social framing goes together with an affective framing. Johnson and Grossmann argue for the existence of a relatively independent network of brain areas for processing social stimuli, in short, a "social brain". This network comprises various areas that participate in processes such as face processing (including eye gaze processing), perception of emotions (in faces and in speech), perception of biological motion, human goal-directed action, and joint attention (Grossmann & Johnson, 2007). The level of neural processing drives cognitive processes which, under the additional influence of motivational and social factors, lead to social behavior and experience. Together, these three levels constitute the new area of "social cognitive neuroscience" (Lieberman, 2006). "Social developmental cognitive neuroscience" then takes into account the developmental aspect (Zelazo, Chandler, & Crone, 2009). Yet, an explicit mentioning of affect in (developmental) cognitive neuroscience is absent, still. Hence, there is a clear need in the future to add affect to the present view. In which way this could be done I want to exemplify with a recent proposal made by Kuhl (2007) on the role of the social brain in language acquisition.

Kuhl (2007) interprets the role of the "social brain" in terms of "gating" of the more computational aspects of language acquisition (phoneme discrimination, word segmentation, referential lexical learning) in the first year of life and beyond. Without social interaction language understanding and acquisition is impossible. In their own language learning studies Kuhl and her colleagues present evidence that 9-month old American infants learned to discriminate phonemes of a foreign language, namely Chinese, only through live interaction with a real person who played with them and not from TV or from auditory presentation alone (Kuhl, Tsao, & Liu, 2003). This phonemic learning was long-lasting and still showed up even after 33 days following the 4-5 week-long exposure to Chinese. In another experiment by Goldstein, King, and West (2003) the role of mothers' contingency on their infants' vocalization was tested. In a contingent condition mothers reacted responsively with smiling, approaching, and touching their infants when they vocalized. In a non-contingent control condition the same mothers' positive responses were shown to the child on a TV monitor – however, with a temporal delay that corrupted the original contingency. As a result, infants with contingent mothers showed more and more mature vocalizations (cf. Kuhl, 2007). Kuhl presents two possible mechanisms underlying the impact of social interaction on language development. On the first account, which she calls "attention and arousal" account and which I call here tendentiously "affective" account it is a global mechanism of motivation, attention, and arousal that positively affects language learning. Enhanced attention and arousal may lead to higher quantities and quality of language input which infants subsequently encode and remember better. Infant- or child-directed speech (ID, CD speech), or, "motherese", is one aspect of this first account. Normal children but not children with autism spectrum disorder (ASD) tend to prefer ID/CD over non-ID/CD (Kuhl, 2007, p. 116). ID/CD is phonetically characterized by exaggerated phonemic contrasts and exaggerated pitch contours as well as clear segmentation clues. It is judged as highly affectionate and sensitive speech. On the second account, which she calls "information" account and which I call here tendentiously "cognitive" account, it is the specific information content present in a natural setting that positively affects language learning. For example, social cues such as eye gaze and pointing to a referent may help in word segmentation and later referential word learning. Through "joint attention" to an object of common interest between parent and infant and contingent behavior of the parent the infant learns to appreciate both the communicative intent of her interlocutor as well as the verbal labels of the objects referred to. This situation is referred to as "triadic interaction". In sum, interactivity and contingency are both relevant to successful communication.

The idea of social *alias* affective cognition as gating computational learning leads to an extension of Kuhl's original "Native Language Magnet" (NLM) model (Kuhl, 1993, 1994), now called "Native Language Magnet-Extended" (NLM-extended, Kuhl, 2007). The NLM is a perceptual effect that distorts the space around a language-specific phonetic prototype, attracting the phonetic processing to its area. The phonetic prototype acts like a magnet in guiding perception reliably to its center. Gradual variations that always exist in perception are attenuated and categorical boundaries are built up with the result that incoming noisy phonetic information is mapped onto stable phonemic categories that represent the language's distinctive phonemes. Thus, when the infant becomes a skilled listener of and neurally committed to the sounds of her own mother tongue by the end of the first year, it is through the attracting effect of the language magnet. The recent extension allows for a gating of this effect through social aspects, some of which can reasonably be interpreted as affective. Interestingly, however, Kuhl never speaks of "affect" or "affective", only of "social". Insofar the contribution of affect in the first "attention and arousal" account remains implicit. This ambiguity is certainly worthwhile being clarified in future research.

PARENT-CHILD INTERACTION

The way the primary caregivers interact with their infants and how infants respond and stimulate their caregivers is essential for infants' development. In an evolutionary perspective, a successful mutual adaptation of caregiver and child has crucial consequences for the reproductive success of the species (Preston & de Waal, 2002). If both parents and infants can be affected by each other's emotional state they may co-regulate each other's state and behavior. Although a transactional perspective, taking into account the mutual interactions and adaptations going on between parents and child, is certainly the best framework for child development (Baldwin, 1995; Preston & de Waal, 2002; Paavola, 2006), the bulk of the literature is on the impact of parents on child development. There is, indeed, a vast literature on the relation between parent-child and in particular mother-child interaction and the general cognitive and language development of young children (Bornstein & Tamis-LeMonda, 1989; Tamis-LeMonda, Bornstein, & Baumwell, 2001; Tamis-LeMonda & Rodriguez, 2008; Landry, 2008; IJzendoorn, Dijkstra, & Bus, 1995). In particular a sensitive or responsive[2] interactional style has been associated with positive outcomes. How is sensitive/responsive behavior towards infants and children characterized? Parental sensitivity in the interaction manifests itself mainly in affectively rich expressive behavior (facial, vocal), enjoyable body contact, and contingent and reciprocal interaction with the infant. As we have seen in the previous paragraph, contingent interaction in a setting of joint attention may modulate the gating of information that is crucial for language development (Kuhl, 2007). Responsive and sensitive interaction is therefore crucial in this respect. Child-directed speech (or "motherese") is just one aspect of parents' affective language behavior towards their children.[3]

Tamis-LeMonda and Rodriguez (2008, p. 2) point out three aspects of parenting that foster children's language development: (i) frequency of *children's engagement in routine language learning opportunities*, such as book reading and story-telling (ii) *quality of parent-child interaction*, in particular sensitivity/responsiveness, and (iii) offering *age-appropriate learning materials*, such as books and toys. Among these three, we shall focus on (ii) in the following. The quality of early parent-child interaction has been shown to be one the best predictors for early and later language learning. Rich and varied adult speech informs children about a wide range of objects and events and, importantly, on the affective attitude that their parents have towards them. Children whose parents respond contingently to their verbal utterances have a head-start into vocabulary development (ibid., p. 3).

Parents' Sensitivity to Informational and Affective Content in Their Infants' Vocalizations

However, already long before the occurrence of discernable verbal utterances, parents respond intuitively to infants' vocal signals expressing discomfort, comfort, neutral mood, or joy. Parents' affective response to these early vocalizations is considered as the basis for a secure attachment and affective attunement between the interactants (Papoušek, 1992, p. 239). Components of parental responsiveness are (i) the ability to decode information from infants' signals, (ii) readiness to respond (in terms of latency), and (iii) quality of response. With respect to decodability of information, Papoušek (1989) presented various groups of adults with 50 infant sounds (discomfort, comfort, joy) and asked them to estimate the quality and intensity of the sounds. As a result, she found that "voiced sounds in the infant's presyllabic vocal repertoire effectively transmit both discrete information pertaining to the categorical distinction between comfort and discomfort and graded information pertaining to the relative intensity of affective arousal." (Papoušek, 1992, p. 241) Moreover, adults spontaneously attribute meaning to these early vocalizations. With this inclination they anticipate later emerging meaningful and language-based communication with children. Parents' early and mostly correct intuitions about their offspring's vocalizations are an important driving force for building up successful interactions and for language development. Parents' responses are affect-based as well as content-based. Identifying the content and intensity of their infants' affective vocalizations and responding to them affectively results in affective attunement within the dyad (Friend, 1985; cf. Papoušek, 1992, p. 243). Adults' responsiveness, according to Papoušek, depends on perceptual and behavioral predispositions and not on experience, gender, and age. It has the effect of "scaffolding" a dyadic communicative setting within which the infant's vocalization has its proper structural slot already. Any vocal material the infant utters will wind up in this slot and will be interpreted by the parent as meaningful and communicatively successful "speech act". Thus, infants find themselves already in interactional settings. Depending on their cognitive and language development, they grow into this framework and contribute increasingly more explicit, language-based utterances to the communication.

The Role of Prosodic Contours in Early Language Development

The modulation of speech that adults display in their speech – be it child-directed or not – does not only carry affective information that is readily exploited by the human mirror neuron system (see appendix) but also structural information about the prosodic properties and parts of speech of the native language. In this respect, it is particularly the melodic contours that carry such information. Prosodic contours help to guide infants' attention to basic units of speech (Papoušek, 1992, p. 245). They are thus in the service of parsing. In fact, post-natally and even pre-natally, infants are sensitive to the prosodic information in the speech signal. Thus, Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, and Amiel-Tison (1988) could show that French neonates can already distinguish their native language French from a foreign language that belongs to a different rhythmic class, such as Russian. Further studies have confirmed this early discriminative ability with other languages (Nazzi, Bertoncini, & Mehler, 1998; Nazzi & Ramus, 2003). What makes prosody of major importance in early language acquisition is the fact that it provides cues for segmenting the seamless speech stream into units, in particular words. Later, at around the age of one year, when semantics comes in, these candidate units can be mapped onto meaning and the first words, which are pairs of form and meaning, emerge. The languages of the world can choose from a variety of prosodic properties on which segmentation may rely, such as word stress (English, Dutch, German), syllable structure (French), syllable weight (Japanese), and vowel harmony (Finnish, Turkish) (Cutler, 1999; Cutler, Norris, & McQueen, 1996). Between 6 and 10 months of age, infants set the rhythmic parameters of their native language with which they will subsequently parse any language input, native or foreign.

The Impact of Affective Mother-Child Interaction on Early Communicative and Language Development

Studying the impact of parent-child interaction on early communicative and language development requires good instruments to assess quantitative measures in both areas longitudinally. Paavola (2006) conducted a study on the relation of early mother-child interaction on preverbal communicative behavior and later language behavior of 27 normally developing Norwegian infants and their mothers. In this study, a multitude of behavioral inventories are applied[4]. Among these the CARE (Child-Adult-RElation)-Index (Crittenden, 2004, 1988) is an assessment tool of adult-child interaction that –unlike other attachment measures – can be used from birth onwards. It measures qualitatively and quantitatively various interaction styles of the adult (sensitive, controlling, unresponsive) and the child (cooperative, compulsive/compliant, difficult, passive) on 7 expressive scales: (1) facial, (2) vocal, (3) position and body contact, (4) affective, (5) contingent, (6) controlling, and (7) choice of activity. The main construct of the CARE-Index is sensitivity in play, which is defined as "any pattern of behavior that pleases the infant and increases the infants' comfort and attentiveness and reduces its distress and disengagement." (Crittenden, 2004, p. 6) Paavola (2006) takes a "transactional view" on the communicative dyad, that is, not only parental behavior is important for infants' communicative and language acquisition but also the child contributes to its success. Moreover, the child's communicative and linguistic behavior influences the adults' behavior such that a framework of mutual transaction is established within which positive feedback enhances and negative feedback attenuates the acquisition process. In the beginning sensitive mothers show some degree of directive control also, as exemplified in initiation of communicative acts and elicitation of conversation. In the course of the development, however, the role of the child becomes increasingly important for the developmental outcome: "It is possible that linguistic development becomes increasingly child-driven over time (…)". (Paavola, 2006, p. 77)

Paavola (2006) found that a high degree of sensitivity of the mother was related to a high amount of communication with the child, in terms of number of communication acts and number of verbal responses, i.e., highly sensitive mothers communicated most with their children, medium sensitive mothers somewhat less and unresponsive mothers least. Likewise, high and medium cooperative infants uttered more intentional communicative acts than uncooperative infants. Maternal responses and child intentional communicative acts with 10 months in the MacArthur Communicative Development Inventories (MCDI) could predict the numbers of phrases understood and vocabulary comprehension at 30 months. The child measures could also predict vocabulary production, whereas maternal scores could not. Sensitive mothers were able to fine-tune their communication in that they talked about age-appropriate here-and-now topics rather than talking in displaced speech. They elicited more communication from their infants, their speech acts were more descriptive and less directive. Highly sensitive mothers made the most compliments. The degree of sensitivity was not crucial in the upper and middle range of sensitivity, i.e., infants of highly and also medium sensitive mothers showed equally good performance. It seems that "good enough maternal care" (Paavola, 2006, p. 78) was sufficient for the linguistic thriving of their children. A crucial part of maternal sensitivity is an active role and initiation of interactions as well as elicitation of conversation with the infants.

The Impact of Mother-Child Interaction on Phonemic Discrimination

It is well-known that infants start out with the ability to discriminate any phonemic contrast of the languages of the world until the age of 6 months, however, lose it by 10-12 months of age (Werker & Tees, 1984; Werker, 1993; Pater, Stager, & Werker, 2004; Kuhl, 2004). By the end of the first year they have become competent listeners of their ambient language and neurally committed to processing it (Kuhl, 2000, 2004, 2007). However, it is not known whether, besides language-related factors other, extraneous factors may also exert an influence on the timing of this development. Typically, in group studies, the majority of infants show discrimination at 6 months of age, however, a minority does not show it. Likewise, at 10 months of age, the group on average shows no discrimination anymore, however, a minority still does. In attempt to explain (part of) this inter-individual variation, the impact of mother-child interaction (as measured by the CARE-Index) on early phonemic discrimination in 6- and 10-month old infants was studied (Elsabbagh, Hohenberger, van Herwegen, Campos, Serres, de Schoenen, Aschersleben, & Karmiloff-Smith (submitted). The study was carried out in three European labs, in London, Paris, and Munich, with mono-lingual infants acquiring English, French and German. In the native condition of the language task, our subjects had to discriminate between two syllables, [ba] and [da] (both syllables had been recorded in French, English, and German versions, respectively). In the non-native condition they had to discriminate between [da] and a non-native Hindi [da] (with a retroflex [d]). As reported in Elsabbagh et al. (submitted), we found that overall, 6-month old infants discriminated the native as well as between the non-native contrasts, whereas 10-months olds only discriminated the native contrast but not the non-native one anymore. This result is expected on the background of the vast literature on early sensitivity to phonemic contrasts in infants up to the age of 8-10 months after which only native contrasts are still discriminated (Werker & Tees, 1984; Kuhl, 2004, among many others). When we looked at the relation of phonemic discrimination with mother-child interaction, we found that at 6 months of age, it was the group of infants of low-contingent mothers (low on maternal sensitivity) that discriminated the non-native contrast whereas infants of high-contingent mothers did not. From this finding we concluded that 6-month old infants of highly contingent mothers had already proceeded so far in their language development that they had already lost their original universal discrimination abilities. In this case, absent discrimination ability is a sign of a more mature language development. We explained this higher maturity of the infants of sensitive mothers in terms of (i) more language input the infants had already received (see Paavola, 2006) and, more importantly, in terms of (ii) contingent interaction. Contingent interaction is also characterized by sensitive verbal and non-verbal interaction. Contingently responsive mothers provide more mutual gaze, turn-taking (verbal and non-verbal) and mutual affect, which may foster the development of communication and language as opposed to the non-contingent behavior.

EMOTIONAL EXPRESSION IN LANGUAGE

A particular area of research is devoted to the relation between language and emotional expression in language, as in affective prosody, and how this relation develops from early childhood onwards.

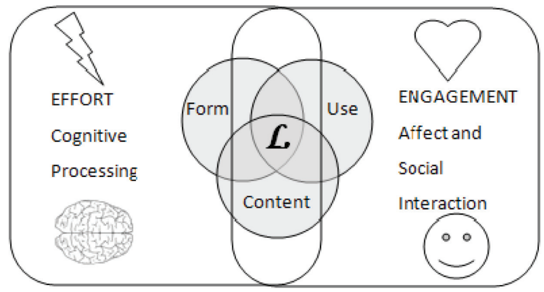
Bloom's Intentionality Model of Language Acquisition

In taking a functional perspective, Bloom (1998) sees "the convergence of emotion, cognition, and social connectedness to other persons" (p. 1272) at the heart of language development in the second year of life. In her view, children learn to represent the contents of their mind (belief, desires, thoughts) in language for the purpose of relating to other persons, and to communicate to them their own thoughts and feelings. These two aspects, (i) representation of conscious mental contents and (ii) communication, that is, sharing of these conscious mental contents with others, constitute Bloom's and Tinker's (2001) "intentionality" model of language acquisition. Her interactive model of language development rests on the two concepts of "engagement" and "effort". The first concept, "engagement", relates to the child's affective and social directedness and interactions with her communicative partners, that is, the child is engaged into the social world *via* language. The second concept, "effort", relates to the resources she has to spend on the cognitive processing of language, that is, the child has to work out and compute the threefold relations that exist within language. Language itself comprises form (sound, words, and syntax), content (meaning), and use (pragmatics). Figure 1 depicts the two intersecting spheres of effort and engagement, at the core of which is language (L), embedded itself in its three components (form, content, and use).

Effort and engagement stand in a relation of fruitful tension and may, at times, dominate the development more or less. Three principles that relate to emotion, social development, and cognition, govern this dialectic tension, (i) relevance, (ii) discrepancy, and (iii) elaboration (cf. Bloom, 1998, p. 1273):

i. *Emotion and relevance*: According to Bloom, children's language learning is tied to their emotional life, which they want to share with others. They communicate to their social partners what is relevant to them and thereby engage themselves into the social world.
ii. *Social development and discrepancy*: A child's social development is dependent on her being understood by her caregivers. If her language and other communicative expressions are insufficient for conveying her intentional states, the child has to acquire the necessary expressive linguistic means to do so successfully. The discrepancy between what the child expresses linguistically and what her social partners understand and *vice versa*, what the caregivers express and what the child understands, is the motor of language development. This motor stops only when the child has fully mastered all aspects of language.

Figure 1. Redrawing of Bloom's intentionality model of language development in terms of effort and engagement (1998, p. 1273; cf. also Bloom and Tinker 2001, p. 14). L = Language[5]



iii. *Cognition and elaboration*: As the child's intentional states become increasingly complex, the linguistic means which represent them also have to become more elaborate. In this sense the child has to invest more resources into the expressive and representative power of the linguistic system.

Bloom and her colleagues studied the relation between the expression of affect and of language in children's spontaneous communication from 9 months to 2 1/2 years of age. This period covers three important language milestones: the first words (FW), the vocabulary spurt (VS), and simple sentences. Subjects varied in their onset of these three milestones and also in their expression of positive, negative, and mixed affect. Importantly, the expression of emotion was a function of the age when the three language milestones were mastered. Early word learners were more often in states of neutral affect, whereas late word learners showed increased emotional expression during that time. Taken together, these complementary findings suggest that children invest their mental efforts either into word learning or into emotional expression. Doing both at the same time – learning words and expressing emotions – seems impossible. Rather, there is a trade-off between the language and emotion which results in the two types of learners.

After equating children for their level of language learning and then looking at their emotional development two findings were revealed: (i) during the second year the expression of emotions was constant, only language increased and (ii) children did not express less emotional content when they learned more words, that is, they did not learn to utter words instead of emotional expressions. The words the children learned during that period were basic words and not emotional words. The above-mentioned trade-off between language and emotional expression can be understood in terms of cognitive load and interference. Uttering words and expressing emotions is both effortful and taps a partially shared pool of resources. Bloom and colleagues found that those words that went along with emotional expressions were among those learned first and used most frequently whereas those words that were uttered without emotional expression were among those learned recently and used less frequently. Taking these two observations together, it can be concluded that the effort spent on learning words interferes with the concurrent expression of emotion. Only words that are highly over-learned can afford a concurrent emotional expression. Often, these emotional expressions were positive, indicating that an objective the child had represented with that word had been achieved and no further effortful mental computations had to be carried out anymore.

Looking into the temporal dynamics of the interaction between affective and language expression on a second-to-second basis, Bloom and colleagues found that the emotional expression was built up and preceded the linguistic expression, e.g., a word. Shortly before the peak in emotional expression the word would be uttered and soon after its completion it would fade again. This fine-grained analysis of the time-course of expression of emotion and language is evidence for both components in the acquisition process – effort and engagement. Children speak about relevant things that are on their mind. In the moment of utterance there is mutual adaptation between emotional engagement and cognitive effort. What has been found for word learning is even more evident for sentence learning, the third milestone. Sure enough, uttering whole sentences is even more cognitively demanding than uttering a single word at a time. In addition, the child is now challenged with the temporal synchronization of an emotional expression across the sentence or across parts of it, i.e. with phrasal or sentential prosody. The same pattern as found in word learning repeated itself again, namely children who had learned syntax earlier showed more emotional expression whereas late learners of syntax showed little emotional expression.

While the developmental course of emotional and language expression that Bloom and her colleagues unraveled, is unquestionable and can be nicely explained in terms of the dialectics between cognition and social engagement, their functional approach may be considered more controversial. The functionalist view, according to which the development of language can be explained in terms of its communicative function, i.e., the child's language learning in order to communicate successfully with the environment, has encountered some serious criticism. Among the arguments against functionalism the criticisms of circularity of functional argumentation and lack of "dysfunctionality" is most prominently (for a critical discussion of functionalism in linguistics, see Keller, 1997). Bloom argues that the need for communication or the fact that communication is successful can explain why language has the structure it has. However, there is no way in which the presumed function of language – communication – can explain why a sentence has the form it has. There is simply no causal chain from the purpose of communication to the form of language. This relation can be explained better by other, more formal theories of language. Generative grammar with its focus on the architecture of linguistic representations is a better candidate, in this respect. The answer to the puzzle of linguistic structure may lie in "principles of neural organization that may be even more deeply grounded in physical law (Chomsky, 1965, p. 59, as cited in Chomsky 2007, p. 14). Generally, those physical principles will be principles of computational efficiency. The human cognitive ability of recursion that yields hierarchically structured expressions is a better explanation in this respect than the emotionally motivated wish to communicate more complex thoughts to one's social partners. After all, one first needs to have some mechanism that enables communication before one can praise the benefit of communicating. Such a mechanism can only be a cognitive one, though, which, then, can be co-opted for communicative purposes.

A more promising view on the role of emotion for language acquisition, therefore, is the "gating" view (Kuhl, 2007) which holds that social and emotional factors direct the child's attention to language as a part of the "social brain" (Grossmann & Johnson, 2007, among others). Through this gating mechanism language acquisition is not explained but rather constrained. This view implies that overall the social and emotional relevance of communication is a supportive factor for language learning. However, Bloom's findings rather show that emotional and cognitive expression can also hinder each other on a more local time-scale, since they seem to partially tap into the same resource during on-line production. Therefore, it seems important to distinguish carefully global from more specific aspects of the relation between language and emotion as well as comprehension from production.

Infants' Sensitivity to What is Said Versus How It is Said

While Bloom studied the relation between emotional and linguistic expression in production, other researchers have looked at this relationship in comprehension. A communicative utterance always comprises language and paralanguage, among the latter affective prosody, facial expressions, gestures, etc. In normal discourse, both sources of information are congruent with each other. However, they may also conflict as when language ("what is said") and paralanguage ("how it is said") contradict each other. These special cases allow particularly well to study the relation between language and affective expression.

Early studies by Fernald (1989, 1993) and Lawrence and Fernald (1993) on the effect of conflicting messages in children of various ages showed that 9-month old infants regulated their behavior more in terms of paralanguage whereas 18-month old infants relied more on language. They concluded that in the first year of life "the melody [is] the message".

Following up on these seminal investigations, Friend studied such conflicting messages in children of various ages (2000, 2001, 2003) with the "social referencing paradigm". Social referencing typically takes place in a novel and ambiguous situation where the infant refers back to an interacting human in order to acquire cues regarding how to handle the situation. In the experiments reported here, interaction was with a female character on a video monitor. In Friend (2001) 15-16-month old infants were presented with novel toys which they were encouraged or discouraged to play with by means of language or facial and vocal paralanguage, respectively. The language and paralanguage cues were either congruent (+language/+paralanguage; -language/-paralanguage) or incongruent (+language/-paralanguage; -language/+paralanguage).[6] As expected, as a group, infants relied more on paralanguage, that is, if the facial and vocal affect was disapproving, infants hesitated longer to grasp the toy and played for a shorter time with them, even if the verbal message was approving and *vice versa* for the reverse combination of cues. However, on the individual level reliance on language was positively correlated with lexical acquisition, as measured by the McArthur Communicative Developmental Inventory (CDI), that is, the more words the infants knew the more they relied on language. This study tapped into an important transition phase characterized by ample inter-individual variation in language development as well as in behavior regulation. This transition gives way to a subsequent phase where children rely more heavily on language cues, suppressing any affective information from the paralinguistic channel. This subsequent stage was studied in Friend (2003) with 4-year olds in the same social referencing paradigm. She found a general effect of the lexical content of the verbal message such that positive messages led to faster approach of the toy as compared to negative messages, irrespective of the +/- affective information. The effect was pervasive in the group: 54-71% of the subjects were more regulated by language as opposed to paralanguage, as measured by the dependent variables “play with novel toys” and “delay to approach”, respectively. However, lexical content had a gradual rather than an all-or-nothing effect on behavior and at least in one of the variables, there was a considerable minority being regulated by paralanguage. Summarizing, this study confirmed the general “lexical bias” in the behavior regulation of older children (4 years). Individual analyses, however, point to differential regulation by language and paralanguage on the micro-level.

This line of studies parallels the findings of Bloom and colleagues, as discussed above. Both in perception and production studies the same developmental pattern emerges: infants (by 9 months of age) first attend to information conveyed through the affective channel. When they start to learn words (by 18 months of age), the lexical channel takes over and pushes back the significance of the affective channel. What is interesting is that both channels stay separate for a protracted period of time. Bloom explains this with the effort that computing both kinds of information requires. It has to be done sequentially before it can be done simultaneously. Only at some later stage can children eventually integrate the information on the two channels. In the perceptual studies at 4 years of age the language channel still dominates the affective channel. From these, two important questions arise: 1. Is there more evidence for this developmental sequence and 2. When does the integration eventually take place?

Language and Emotion in a Cross-Modal Perspective

Evidence for the same developmental sequence and for the eventual integration of language and affect in children comes from the cross-modal comparison of the expression of language and emotion in spoken and signed languages (Reilly & Seibert, 2003). Sign language offers a unique perspective on the relation of these two systems since the same means – in particular facial gestures – are used for both systems: syntax and emotion. On the one hand, sign languages, beyond certain manual parameters, make use of facial expressions (eyes, eyebrows, facial and mouth gestures) and body leans in order to express syntactic properties, e.g., sentence types (*yes-no* and *wh*-questions, topic, negation, conditionals, relative clauses), scope of syntactic and/or prosodic domains and signing perspective (Sandler & Lillo-Martin, 2006). Less prominently, these features are also used to convey lexical information, e.g., the manual sign HAPPY has an obligatory happy facial expression, as a part of its lexical content. On the other hand, facial and bodily expressions are used for conveying emotional information such as the signer’s current affective state, subjective attitude towards and evaluation of the present topic. For example, raised eye-brows syntactically indicate a *yes-no* question, whereas, in a different context, affectively they indicate surprise. While the muscles used to innervate these facial expressions are the same, however, the morphology, i.e., the form of these expressions, the context in which they are used, their scope, and their timing characteristics differs in both domains (Reilly & Seibert, 2003). Syntactic prosody is linguistically constrained. It occurs in the context of clauses, is carried out with a higher muscle tension, and has a strictly defined temporal distribution in the sentence with neatly marked onsets and offsets and characteristic prosodic contours. This is because syntactic prosody is a part of the language core computational system. In adults these two systems – language and emotion – are clearly separable but highly integrated in normal, healthy subjects. The language and emotional systems show a characteristic break-down in lesions. Deaf signers who suffered either a right- or left-hemispheric stroke revealed a double dissociation of affective and syntactic expression. A stroke in the right hemisphere leads to flat emotional facial expressions, however, unimpeded syntactic prosody whereas the reverse pattern holds for a left-hemispheric lesion (Poizner, Klima, & Bellugi, 1990; Corina, Bellugi, & Reilly, 1999). Thus, a deaf patient having suffered a right-hemispheric stroke cannot raise her eyebrows anymore as a sign of surprise but still can raise her eyebrows in a syntactic *yes-no* question. How do children acquire and – most interestingly – integrate both? Reilly and Seibert (2003) discuss the two systems in a dynamical developmental framework that takes factors with different scopes in consideration: "…, these systems interact in a dynamical way, according to both specific (i.e., contextual) and potentially universal (i.e., developmental and biological) factors." (p. 553) The emphasis of the remainder of this paragraph will lie on the acquisition of grammatical and emotional facial expressions in American Sign Language (ASL) and how their use develops in narratives.

Acquisition of grammatical and emotional facial expression in ASL: Deaf infants younger than two years of age sign one-word utterances with prosodic expressions, e.g., HAPPY with a happy face or WHAT with furrowed eyebrows. On closer inspection, however, these holophrases appear to be unanalyzed chunks in which the facial expression cannot be detached from the manual sign, indicating that they have not yet been analyzed separately as a systematic part of the sign or phrase. Also, the timing, scope, and individual characteristics of these facial expressions are often erroneous. Around 2;6 years, interestingly, the facial expression disappears. Deaf toddlers then produce grammatical constructions which need syntactic prosody, e.g., *wh*-phrases, with correct manual syntax but with totally blank faces (Reilly & McIntire, 1991). Only at older ages than four years the prosody reappears and becomes integrated with the manual component. The correct morphology, scope and timing characteristics have to be learned in a protracted time period over the next couple of years. Summarizing, a non-linear developmental sequence can be stated in the acquisition of syntactic prosody in deaf ASL-acquiring children: (i) initial co-articulation of prosody alongside manual signs in the sense of unanalyzed chunks (<2 yrs), (ii) absence of syntactic prosody (2;6-4 yrs), and (iii) reappearance and integration (> 4 yrs). Reilly and Seibert interpret this developmental sequence in terms of a reanalysis of affective and syntactic resources feeding it: "Initially, (…), both language and emotional expression have access to a general, underlying symbolic function. Children at the one-sign stage draw on early affective and communicative abilities collaboratively with their first signs. However, as language emerges, especially as syntax develops, and the child begins to combine signs, there is a bifurcation of systems such that language and emotion unfold as independently organized and differentially mediated systems, each following its own developmental path." (p. 547)

Acquisition of narrative skills in hearing and deaf children in terms of emotional expression: In narratives, language and emotion must be integrated in order to convey and understand the narrated episodes. Evaluative elements represent the attitude of the narrators towards the story and the characters in it. These evaluations typically give a narrative its meaning. They can be conveyed in various ways: lexically, syntactically, and paralinguistically, as with affective prosody, facial expressions and gestures. Reilly and Seibert studied hearing and deaf children between 3 and 12 years of age with the famous "Frog, where are you?" story (Bamberg & Reilly, 1996). For the hearing children they found that the youngest age group (3-4 yrs) could not really tell the story well. Their productions were structurally weak, however, were full of vocal prosody, stress, vowel lengthening and showed sing-song like intonation contours. There was little facial expression. The role of the affective prosody was to "glue together" the structurally unconnected parts of the story (Reilly & Seibert 2003, p. 549). The school-children (7-8 yrs) told structurally more elaborate stories, however, almost without any paralinguistic evaluation, flat affect and stereotypical. Bamberg and Reilly (1996) conceive of this developmental pattern so far as "a transition from paralinguistic to linguistically conveyed evaluation, that is, a lexicalization of affective expression that appears to occur during the early school years." (Reilly & Seibert 2003, p. 549) As an explanation for the early prevalence of prosodic means and its later ban the authors propose "that paralinguistic expression functions as a support system and stepping stone into the lexicalized expression of evaluation." (p. 549) Only the 10-11-year olds could reintegrate these two systems again and tell structurally complex and affectively rich stories. Summarizing the developmental process of narratives, the now familiar tripartite sequence re-occurs again: (i) narratives with much affective prosody but little structure (3-4 yrs), (ii) lexicalization of affective evaluation, flat emotional expression (7-8 yrs), (iii) integration of affective paralanguage and structurally elaborate language (> 10 yrs).

Would ASL-acquiring deaf children show the same sequence as their hearing counterparts, or, due to the different modality, a different one? Adult signers heavily use paralinguistic means for evaluative purposes in narrations: affective prosody, emotional facial expressions and gestures. This is a modality-specific phenomenon which stands in marked contrast to adult hearing subjects who rely more on linguistically encoded evaluations. The development of signing children, however, was found to be quite similar to that of their hearing counterparts and all three stages were confirmed, as well. In perspective taking, for example, adult signers make body shifts and facial expressions so as to adopt the position and the facial expression of the story character who they are about to quote directly subsequently. Preschoolers show facial expressions as well, however, their occurrence is incorrectly timed, unspecific for particular characters, and erratic. From 5 years onwards, however, all children use the lexical expression SAY in order to introduce the quotation, which, according to the authors "reflects a linguistic reorganization: direct quotes are introduced lexically before the non-manual behaviors are linguistically integrated with the manually signed quote." (p. 551) Again, only the older children were able to produce the non-manual affective prosody synchronized with the manual lexical signs. This parallel development of narrative skills in hearing and deaf children led the authors to the conclusion that this development was modality-independent. They summarize their findings as follows: "This recurrent pattern – first relying on affective means, then moving on to a manual lexical strategy, and finally integrating both channels such that the facial behaviors are now under linguistic control – appears across structures, from single lexical items to phrasal, clausal, and now discourse level structures. These data provide strong evidence that the child's prelinguistic, emotional abilities are not directly accessible to the linguistic system, even when they continue to convey affective information." (p. 551)

Summarizing the common developmental pattern in the reviewed areas – the role of linguistic and paralinguistic information in early language production and perception, the cross-modal similarities in speaking and signing children's narratives and in the expression of affect in sign language acquisition – there seems to be a recurring universal tripartite sequence of

i. an undifferentiated mixed system dominated by affect over language
ii. a strictly language-dominated system in which affect is lexicalized and language governs behavior
iii. a bifurcation into two mature systems, emotion and language, and a concurrent re-integration of both systems

This development resembles the classical three steps "thesis – anti-thesis – synthesis". It seems to be not only applicable to various domains but also to occur independently of age. Note that in Friend's studies on the role of paralanguage and language on children's behavior regulation these steps occur, roughly, at (i) 9 months, (ii) 15-16 months, and (iii) 4 years. In Reilly and Seibert's study on the expression of emotion in narratives, however, they occur at (i) 3-4 years, (ii) 6-7 years, and (iii) 10-11 years. This décalage suggests that the underlying process is not so much age-dependent, i.e., a matter of (perhaps neuronal) maturation, but one that needs to be recapitulated in different domains at different times. The model that best accommodates these properties is Karmiloff-Smith's *representational redescription* model (1979, 1986, 1992). This model states that in the course of acquiring a skill or cognitive faculty, irrespective of age, subjects entertain different representations as they proceed. Initially, these representations are implicit (e.g., represented sensu-motorically), however, later on, they become more and more explicit, which makes them amenable to linguistic expressions. The earlier levels are not lost; rather they are "redescribed" in a more general, abstract cognitive vocabulary that can be shared by other cognitive systems that can access that level as well. At these higher levels, integration is possible, even with the original "lower"-level format. In the present context, the formal language system having become sufficiently autonomous from the emotional system can be re-integrated with it again. This re-integration seems to have to be achieved for different domains separately, however, in a similar, tripartite, process.

The developmental sequence of an early emotional language phase of young children and the later "ban" on emotional expressions in their narrative productions may also be related to a novel evolutionary account in terms of "self-domestication" of humans in the course of their social-cognitive evolution. In their "emotional reactivity" hypothesis Hare and Tomasello (2005; Hare, 2007) try to explain how humans managed to evolve socially-relevant cognitive abilities that make them uniquely human. This hypothesis holds that the evolution of human social problem solving might have proceeded in two steps: First, the level of emotional reactivity towards con-specifics needed to be lowered, i.e., they had to feel comfortable and be capable of interacting cooperatively in a group; second, and contingent upon the first development, socially relevant cognitive abilities such as action and intention understanding and theory of mind came under selection pressure and evolved. For the development of language abilities a similar two-step process occurred. First, the original, emotion-based call system was abandoned or, more precisely, not selected as a substrate for linguistic computation, i.e., the emotional reactivity was disconnected from our modern language faculty. Once this had happened, the evolutionary pressure acted directly on the cognitive faculties of thought and language until a potent computational language system emerged. In modern humans, thus, emotion and language are clearly different systems, however, systems that still – or again – interact with each other very strongly on various linguistic levels such as prosody and lexical meaning, maybe even syntax. While they are well integrated in adults, the developmental perspective reveals that at some point the two were dissociated – maybe had to become dissociated in order to reach higher levels of performance before becoming reintegrated again in the mature form that we witness today in contemporary adults.

EMOTIONS IN FIRST AND SECOND LANGUAGE LEARNING

Harris et al. (2006; cf. also Caldwell-Harris, 2008) lament that the "emotion revolution" which started in the cognitive science in the 1990'ies has not yet captured psycholinguistics, including first (L1) and second language (L2) learning and bilingualism. They claim that the fact that speakers experience more emotions in their L1 as compared to their L2 is due to the more emotional context of language acquisition and not so much to the time of language acquisition as such. In their "emotional contexts of learning hypothesis" they propose that "words and phrases that are acquired early will have strong connections to the amygdala (LaBar & Phelps, 1998), because early language develops at the same time as emotional regulation systems (Bloom & Beckwith, 1989). Later learned language may have a more purely cortical representation, lacking connections to subcortical areas (Lieberman, 2000)." (Harris et al., 2006, p. 271) The direct connections of words with their associated emotions strengthen their representation and facilitate their processing. As a consequence, an L2 can, in principle, achieve the same "emotionality" if acquired in an equally emotional context as the L1. "Age of acquisition" is not the causal factor, but "emotional context of learning". The L1, on the one hand, is so emotional because it has been acquired in an emotionally rich context of binding to the primary caregivers: "Early age of acquisition thus functions as a proxy for a more emotional context of learning." (Harris et al., 2006, p. 274) Early age and proficiency are systematically confounded with the emotionality of the learning context. The L2, on the other hand, is less emotional because it is acquired later in a large variety of settings, most of which are formal such as learning at school. This hypothesis is compatible with the results of their studies in which they investigated the psycho-physiological reaction, as measured by the skin conductance response (SCR), of subjects belonging to various bilingual learner groups to emotional words. In study 1 they investigated Turkish-English bilinguals; in study 2 they investigated (i) subjects coming to the US as adults, (ii) subjects coming to the US as children, and (iii) bilinguals born in the USA as children of immigrants. The subjects listened to emotional phrases and neutral words in their L1 and L2, respectively. The results showed that age of acquisition of the L2 influenced the emotional response, however, only for late learners who had better language proficiency in their L1. When L2 had become the stronger language, the electro-dermal responses of the subjects to emotional stimuli in L1 and L2 did not differ.

Harris et al.'s hypothesis requires a different model of language processing and of the mental lexicon. In the classical Levelt model of lexical access (Levelt, Roelofs, & Meyer, 1999; Levelt, 1999), words, are related with abstract conceptual structures which then may have connections to some emotions. The "emotional contexts of learning" hypothesis, however, suggests a direct connection between a word and its associated emotion as well as with the context in which it was learned and is predominantly used. This view is in favor of the connectionist idea of a distributed associative lexicon. Their hypothesis also requires a different language learning model, namely one that is age-independent and which poses the explanatory burden pre-dominantly on the emotional context of learning.

Despite the intriguing arguments from the proponents of the emotional contexts of learning hypothesis, there are counter-arguments against several of their claims. If it were true that it is not the age of acquisition but the emotional context of acquisition that decides on the success of the learning process and the emotionality in the achieved language, then there should also be a non-negligible number of low-proficiency L1 adult speakers who, as children, did not have the luck of having been raised in an ideal emotionally supportive and rich environment. However, also such individuals acquire their native language successfully. Furthermore, there is evidence that L2 learners regularly switch back to their native language in the number domain, e.g., when it comes to counting and carrying out mathematical operations in real-time. However, arithmetic is certainly the least emotionally dependent domain. Why, then, should subjects fall back on their L1 in an area that is unrelated to emotions? Lastly, there is good evidence that a critical period for language learning exists, which is denied by Caldwell-Harris (2008). Even connectionists, on who these authors rely when they argue for distributed lexical representations, nowadays model brain plasticity, that is, the timing of the developmental susceptibility for learning in various sensory domains (Elman et al., 1996). Unless these and possibly other inconsistencies are resolved, the emotional contexts of learning hypothesis is nothing more than – a hypothesis.

DELIMITATION OF LANGUAGE, COGNITION AND EMOTION FROM AN EVOLUTIONARY VIEW

In the traditional view, emotion and cognition (comprising language), have been strictly separated. However, more recent approaches propose a very tight integration, if not inseparability of language, cognition, and emotion (at least for adult processing), as most strongly expressed by Pessoa (2008) and Caldwell-Harris (2008), and of direct mapping between these domains, as in the mirror neuron and embodiment literature (Rizzolatti & Craighero, 2004; Foroni & Semin, 2009, among many others). However, whether we espouse a separatist view or an integrational, embodied view, may depend on the level of our analysis. As Ochsner and Barrett (2001) point out, a separation of cognition and emotion seems to make sense for researchers that study high-level phenomena, such as phenomenal experience or behavior, e.g., there are distinct feelings connected with the experience of emotional or neural processing. This clear separation cannot be maintained, however, at lower levels of analyses such as information processing and neural processing. As an example, Ochsner and Barrett refer to the evaluative/monitoring function of the anterior cingulate cortex (ACC). This brain structure is implied in the decision whether a current action should be continued or changed. Whether this action is a purely cognitive one, such as retrieving a word meaning, or an emotional one, such as evading a predator, makes the same functioning either cognitive or emotional. The activation of the ACC does not tell us straightforwardly what kind of computation is carried out, in contrast to, e.g., the amygdala, whose activation is a better sign of affective computation (in this case, fear). Given this ambiguity at least at the lower levels of information and neural processing, it is not at all clear which criteria should be used in order to distinguish between cognition and emotion: experiential, behavioral, computational, or neural criteria? (Ochsner & Barrett, 2001, p. 65) Ochsner and Barrett ask critically whether only processes that are accompanied by conscious experiencing of feelings should be called emotional or also those that only show the involvement of a brain area known to be involved in affective computation despite the absence of any experiential emotional state. Hence, would a measured level of arousal be sufficient to define a state as affective? This ambiguity is far from being resolved. It is important, however, to frame it in a theoretical approach that has the capacity of integrating the various domains. It seems as if the emerging discipline of "social cognitive neuroscience" is the most suitable and promising framework for this task (Ochsner & Barrett, 2001; Lieberman, 2007). If, as has been attempted here, a developmental perspective is taken, then the framework is called "developmental social cognitive neuroscience" (Zelazo, Chandler, & Crone, 2009). These two novel research areas certainly have a very promising future.

The paradigm of embodied cognition that was initially put forward to explain action and intention understanding, memory, and language in general, has recently also been used to explain the impact of emotion in first and second language acquisition. The claim of embodied cognition is that emotional words are acquired in emotional contexts (Harris et al., 2006) and that these contexts are (partially) relived or simulated somatically, as in motor resonance, when we process them (Foroni & Semin, 2009). However, it is not clear at all whether motor resonance and simulation contribute to lexical processing. As argued above, conceptual and lexical representations must be clearly demarcated. The proposed grounding seems to take place at the conceptual level where numerous multi-modal systems interface with each other and jointly contribute to the pre-verbal conceptual representation of a word, among them the motor system. In this respect, the exact time-course of the processing of these various sources of information needs to be investigated in more detail. In Foroni and Semin (2009) motor resonance sets in only after 250 ms of stimulus presentation and unfolds at least during a period of 2 seconds. Such a long dwell-time for a single word is not viable in normal language processing where up to four syllables are processed within one second and where simply no time is available for such extended reverberation processes. It is more likely that motor resonance is an optional post-lexical process that may take place when the conceptual structure is accessed in the later course of lexical access of the perceived word. It is an open question whether it can then still contribute to word comprehension. This scenario does not preclude the attested role of motor resonance in influencing later affective evaluations and prompting actions. However, Gallese (2008) counters the criticism against this "late motor imagery hypothesis." (p. 325) He adduces findings from an ERP-study on single word processing (Pulvermüller, Härle, & Hummel, 2000) that detects modality-specific activation of somatotopic brain areas after a delay of ~200ms after word onset. He evaluates this delay as short enough to support lexical, and not just post-lexical, processing. Niedenthal (2007, p. 1005) shares this criticism. She argues that the reaction times found in the brain's modality-specific systems are on an appropriate time-scale whereas older accounts that argued with activation of muscles and viscera may not. If the output of these more subtle brain-internal computations need not have to surface at the level of overt behavior, but may remain confined to the neural interface-level of pre-motor representations, these may still count as grounded, embodied representations. The view that the pre-motor cortex, which supports action as well as language representations, is the site of simulation whereas the motor cortex is the site of execution may indeed be acceptable to both views – the classical cognitivist and the recent embodied view. If the pre-motor cortex is an interface, indeed, it may be called "cognitive" as well as "embodied" – it depends on which side of the interface one is looking at. The theoretical discussion has actually already proceeded beyond a strict dichotomy between these paradigms. Thus, Dove (2009) pleads for "representational pluralism". It may neither be feasible nor advisable to try to ground every concept in perception. For an abstract notion such as "democracy" it may not be possible to show how the conceptual and lexical processing is causally determined by simulation and re-enactment of whatever situations people associate with it.

The grounding of language in the body and in emotional states is still a relatively recent proposal whose viability cannot be evaluated conclusively at this point. However, it touches on some very basic and important topics in the evolution of human cognition and language as well as in language development. For the evolution of language, we know that affective vocalizations, which are also abundant in the animal kingdom, are supported by sub-cortical brain regions, whereas human language is processed pre-dominantly in the left hemisphere of the neo-cortex. Ancient affective cries and calls, though vocally based, have not evolved into modern human language. This is because they are under involuntary, sub-cortical control. If a predator is recognized, the individual must signal the threat to its con-specifics as fast as possible. Language, however, needs to be under voluntary, neo-cortical control. Forming a non-verbal message and communicating it to others is a deliberate act of the mind. Therefore, spoken language presumably evolved from the gestural system of sign language (Corballis, 2002, 2010; Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004). This is because the hands (also the fore-limbs, head and torso), are under voluntary control. The first language thus was a sign language. In terms of Hare and Tomasello's "emotional reactivity hypothesis" (Hare, 2007; Hare & Tomasello, 2005), this step in the evolution of language would then be comparable to the selection of a cognitive ability once the emotional reactivity had been down-regulated. To be more precise, in this case, the whole communication system was shifted to another brain circuitry: from sub-cortical to cortical control. The transition from sign to spoken language was a later process that was presumably supported by the co-articulation of sounds along with hand and mouth gestures (Rizzolatti & Arbib, 1998; Arbib, 2005). Once the imitative capabilities to reproduce the sounds of the concurrent hand or mouth actions without actually having to carry out the actions had become potent enough, language shifted from gestural to spoken language. Hence, the crucial achievement was to free the concurrent mouth gesture from its grounding in the body such that it could become a symbol in an emerging abstract linguistic system. This achievement cannot be appreciated high enough. Insofar the discussion of embodied cognition is misconceived. It is well taken that eventually language is rooted in the action system, however, isn't it more significant that it emancipated itself from it? Actually, proponents of embodied cognition as well as of traditional linguistic theories could be equally content if the significance of both systems, the body and the mind, could be acknowledged with equal rights. Non-reductive philosophical theories of supervenience or emergence seem best suited to mediate in this dispute (Kim, 2006; Stephan, 2006). That emotional prosody, gesturing, and facial expressions, are part of skillful speaking and signing, is owed to the reintegration of the discrete digital linguistic system and the analog emotional system, after their separation. Language, according to Goldin-Meadow (2003), has to accommodate both functions: the systemic and the imagistic function. The imagistic function is what Reilly and Seibert (2003) refer to as the evaluative, affective stance taken in narratives. In spoken languages this division is labor is usually accomplished across the modalities such that gestures, facial expressions and other paralinguistic devices are relegated to the gestural modality while the systemic function is fully covered by the auditory modality (except prosody). In sign languages, the gestural modality can accomplish both functions: the systemic function we call "sign language" and the imagistic function we call "gestures" or "pantomimes". As we have seen, during acquisition, deaf children acquiring a sign language have to understand this division in spite of a common gestural channel in which they are expressed. This functional divergence is often accomplished through temporal abandoning of one function, namely the emotional gestural function, while the linguistic system achieves dominance, before it becomes reintegrated (Reilly & Seibert, 2003).[7] The ability of "stripping off" emotion from language by lexicalizing them is actually an important achievement in human evolution and in human development. In fact, the literature reviewed here hints at such a process (Friend, 2001, 2003; Reilly & Seibert, 2003). Counter-intuitively, the language of school children appears less grounded and less emotional than the language of adults! This surprising finding is not in line with a recent trend in cognitive science to stress the mutual penetration between emotion and cognition but is well in line with Hare and Tomasello's "emotional reactivity" hypothesis which proposes that an emotional down-regulation or self-domestication must have happened some time during the evolution of humans which paved the way for a following selection of genuine cognitive abilities that boosted human's unique cognitive, social, and cultural abilities – perhaps also language.

Speaking of culture, the intimate relation between language and emotion is most clearly expressed in the area of art. Poetry and music are powerful forms of language use that humans in all cultures and times enjoy(ed) producing and perceiving. While the fashions are changing over time, the bond between language and emotion remains constant. Singing children's songs and listening to children's rhymes and poetry probably belong to the most precious memories of any adult person. A baby falling asleep to the soothing sound of a lullaby requires no less than a human brain that is exquisitely adapted to the processing of language and emotion, embedded in a body capable of resonating with the conveyed emotional state.

CONCLUSION

After reviewing a broad range of areas where language, cognition, and emotion interact, the main conclusion to be drawn is that these different systems interact strongly, at times to a degree that they cannot reasonably be separated anymore. However, when a developmental (and evolutionary) perspective is taken, their tight integration appears as the result of a characteristic tripartite sequence of consecutive steps: (i) an undifferentiated affect-dominated system governs the infant's and child's behavior, (ii) a cognitive and language-dominated system emerges in the course of which emotions are lexicalized, behavior is mainly driven by explicit language, and affect expression in the linguistic behavior is strongly attenuated, (iii) emotional expression is re-integrated into language and cognition and the characteristic tight and skillful coupling of both systems as in adult behavior is observed. The distinct cognitive and linguistic system in (ii) appears to have evolved either from the undifferentiated emotion-dominated system in (i) or has selected a different system for its evolution, despite its roots in or remaining connections with the former emotion system. As it became more and more independent, it offered the possibility of re-organization of the affective system as well and allowed for a division of labor between the two. At the same time, it offered the opportunity of re-integration of the two systems. As far as the evolution of this tripartite process is concerned, many aspects of it still remain elusive, however, as far as ontogeny is concerned, the picture is much clearer, as the process can be studied empirically. Besides this substantive conclusion there is another, methodological conclusion, to be drawn here, which perhaps is of even higher significance, namely that of taking a developmental perspective. Only such a dynamical perspective reveals the strands of adult behavior, their (speculative) phylogeny in evolution and their ontogeny in infants and children. We have seen that what looks like inseparable and fully intertwined processes might have developed from an undifferentiated domain to two differentiated and then again re-integrated domains or has chosen a different neural substrate and developmental pathway as in the shift from the vocal call system to sign and then spoken language. Developmental processes taking place in these domains can be looked at neuro-scientifically and/or behaviorally, within one modality (speech) and across modalities (sign). A review of these areas and evaluation of the implications for our understanding of how language, cognition, and emotion interact in the human mind, brain, and body was presented in this chapter. However, the complexity of the subject matter prohibits drawing any pre-mature conclusions; hence at this stage the proposed tripartite scenario aims at stimulating future discussion and research.

REFERENCES

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *The Behavioral and Brain Sciences*, *28*, 105–167. doi:10.1017/S0140525X05000038

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS ONE*, *5*(1), e8759. .doi:10.1371/journal.pone.0008759

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In Moore, C., & Dunham, P. J. (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, New Jersey: Erlbaum.

Bamberg, M., & Reilly, J. S. (1996). Emotion, narrative and affect. In Slobin, D. I., Gerhardt, J., Kyratzis, A., & Guo, J. (Eds.), *Social interaction, social context and language: Essays in honor of Susan Ervin-Tripp* (pp. 329–341). Norwood, New Jersey: Erlbaum.

Bastiaansen, J. A. C. J., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B*, *364*, 2391–2404. doi:10.1098/rstb.2009.0058

Bloom, L. (1998). Language development and emotional expression. *Pediatrics*, *102*(5), 1272–1277.

Bloom, L., & Beckwith, R. (1989). Talking with feeling: Integrating affective and linguistic expression in early language development. *Cognition and Emotion*, *3*, 315–342. doi:10.1080/02699938908412711

Bloom, L., & Tinker, E. (2001). The intentionality model and language acquisition: Engagement, effort, and the essential tension in development. *Monographs of the Society for Research in Child Development*, *66*(4), i–viii, 1–91.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. In Bornstein, M. H. (Ed.), *Maternal responsiveness: Characteristics and consequences. New directions for child development, no. 43* (pp. 49–61). San Francisco: Jossey-Bass.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science*, *5*(1), 72–83. doi:10.1111/1467-7687.00211

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. doi:10.1016/j.tics.2005.08.007

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, *2*(1), 1–19. doi:10.1017/S0305000900000866

Bruner, J. S. (1983). *Child's talk: Learning to use language*. New York: Norton.

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences*, *21*(7), 273–313. doi:10.1016/S0166-2236(97)01214-9

Caldwell-Harris, C. L. (2008). Language research needs an "emotion revolution" AND distributed models of the lexicon. *Bilingualism: Language and Cognition*, *11*(2), 169–171. doi:10.1017/S1366728908003301

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5497–5502. doi:10.1073/pnas.0935845100

Cato, M. A., Crosson, B., Gokcay, D., Soltysik, D., Wierenga, C., & Gopinath, K. (2004). Processing Words with Emotional Connotation: An fMRI Study of Time Course and Laterality in Rostral Frontal and Retrosplenial Cortices. *Journal of Cognitive Neuroscience*, *16*(2), 167–177. doi:10.1162/089892904322984481

Chakrabarti, B., Bullmore, E., & Baron-Cohen, S. (2006). Empathizing with basic emotions: Common and discrete neural substrates. *Social Neuroscience*, *1*(3-4), 364–384. doi:10.1080/17470910601041317

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

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

Chomsky, N. (2007). Biolinguistic explorations: Design, development, evolution. *Journal of Philosophical Studies*, *15*(1), 1–21. doi:10.1080/09672550601143078

Corballis, M. C. (2002). *From hand to mouth. The origins of language*. Princeton, Oxford: Princeton University Press.

Corina, D. P., Bellugi, U., & Reilly, J. (1999). Neuropsychological studies in linguistic and affective facial expressions in deaf signers. *Language and Speech*, *42*(2-3), 307–331. doi:10.1177/00238309990420020801

Crittenden, P. M. (1988). Relationships at risk. In Belsky, J., & Nezworski, T. (Eds.), *Clinical implications of attachment* (pp. 136–174). Hillsdale, NJ: Erlbaum.

Crittenden, P. M. (2004). *CARE-Index: Coding Manual*. Unpublished manuscript, Miami, FL.

Cutler, A. (1999). Prosodic structure and word recognition. In Friederici, A. D. (Ed.), *Language comprehension: A biological perspective* (pp. 41–70). Berlin, Heidelberg: Springer.

Cutler, A., Norris, D. G., &McQueen, Ja. M. (1996). Lexical access in continuous speech: Language-specific realisations of a universal model. In T. Otake and A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies* (pp. 227-242). Berlin: Mouton de Gruyter.

Davidson, R. J. (2000). Cognitive neuroscience needs affective neuroscience (and vice versa). *Brain and Cognition*, *42*(1), 89–92. doi:10.1006/brcg.1999.1170

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, *3*(1), 11–21. doi:10.1016/S1364-6613(98)01265-0

De Gelder, B., & Vroomen, J. (2000). Perceiving emotion by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. doi:10.1080/026999300378824

De Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, *10*(10), 435–441. doi:10.1016/j.tics.2006.08.008

Decety, J., & Ickes, W. (2009). *The social neuroscience of empathy*. Cambridge: MIT Press.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*(2), 71–100. doi:10.1177/1534582304267187

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science*, *15*(2), 54–58. doi:10.1111/j.0963-7214.2006.00406.x

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*, 412–431. doi:10.1016/j.cognition.2008.11.016

Elman, J., & Karmiloff-Smith, A. Bates, Elizabeth, Johnson, Mark, Parisi, Domenico, & Plunkett, Kim (1996). *Rethinking innateness. A connectionist perspective on development*. Cambridge, MA: MIT Press.

Elsabbagh, M., Hohenberger, A., Van Herwegen, J., Campos, R., Serres, J., de Schoenen, S., Aschersleben, G., & Karmiloff-Smith, A. (submitted). *Narrowing perceptual sensitivity to the native language in infancy: exogenous influences on developmental timing.*

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878–879. doi:10.1038/nn1729

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Development*, *60*(6), 1497–1510. doi:10.2307/1130938

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, *64*(3), 657–674. doi:10.2307/1131209

Fodor, J. A. (1975). *The language of thought*. New York: Crowell.

Fogassi, L., & Ferrari, Pier F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, *16*(3), 136–141. doi:10.1111/j.1467-8721.2007.00491.x

Forgas, J. P. (2008). Affect and cognition. *Perspectives on Psychological Science*, *3*(2), 94–101. doi:10.1111/j.1745-6916.2008.00067.x

Foroni, F., & Semin, G. R. (2009). Language that puts you in touch with your bodily feelings. The multimodal responsiveness of affective expressions. *Psychological Science*, *20*(8), 974–980. doi:10.1111/j.1467-9280.2009.02400.x

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science*, *3*, 148–162. doi:10.1111/1467-7687.00108

Friend, M. (2001). The transition from affective to linguistic meaning. *First Language*, *21*, 219–243. doi:10.1177/014272370102106302

Friend, M. (2003). What should I do? Behavior regulation by language and paralanguage in early childhood. *Journal of Cognition and Development*, *4*(2), 162–183. doi:10.1207/S15327647JCD0402_02

Gallese, V. (2001). The "Shared Manifold" hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies*, *8*(5-7), 33–50.

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience*, *3*(3-4), 317–333.

Gallese, V., Fadiga, L., & Fogassi, L., & Rizzolatti, Giacomo. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. doi:10.1093/brain/119.2.593

Gallese, V., & Keysers, C., & Rizzolatti, Giacomo. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403. doi:10.1016/j.tics.2004.07.002

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology*, *16*(18), 1824–1829. doi:10.1016/j.cub.2006.07.072

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1–26. doi:10.1080/01638539909545067

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: a comparison of high dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*(3), 379–401. doi:10.1006/jmla.2000.2714

Goldin-Meadow, S. (2003). *The resilience of language. What gesture creation in deaf children can tell us about how all children learn language*. New York, Hove: Psychology Press (Taylor and Francis).

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *The European Journal of Neuroscience*, *25*(4), 909–919. doi:10.1111/j.1460-9568.2007.05379.x

Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport*, *16*(16), 1825–1828. doi:10.1097/01.wnr.0000185964.34336.b1

Grossmann, T., Striano, T., & Friederici, A. D. (2006). Crossmodal integration of emotional information from face and voice in the infant brain. *Developmental Science*, *9*(3), 309–315. doi:10.1111/j.1467-7687.2006.00494.x

Hare, B. (2007). From non-human to human mind. What changed and why? *Current Directions in Psychological Science*, *16*(2), 60–64. doi:10.1111/j.1467-8721.2007.00476.x

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, *9*(9), 439–444. doi:10.1016/j.tics.2005.07.003

Harris, C. L., Berko Gleason, J., & Aycicegi, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In Pavlenko, A. (Ed.), *Bilingual minds. Emotional experience, expression and representation* (pp. 257–283). Clevedon: Multilingual Matters.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2005). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307. doi:10.1016/S0896-6273(03)00838-9

Hohenberger, A. (2004). S.Goldin-Meadow (2003). The resilience of language. What gesture creation in deaf children can tell us about how all children learn language. New York, Hove.: Psychology Press (Taylor and Francis). *Linguist List 15-683*.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, *3*(3), 529–535. doi:10.1371/journal.pbio.0030079

Kaplan, J. T., & Iacoboni, M. (2006). Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social Neuroscience*, *1*(3-4), 175–183. doi:10.1080/17470910600985605

Karmiloff-Smith, A. (1979). Micro- and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive Science*, *3*(2), 91–118. doi:10.1207/s15516709cog0302_1

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, *23*, 95–147. doi:10.1016/0010-0277(86)90040-5

Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Keller, R. (1997). In what sense can explanations of language change be functional? In Gvozdanovic, J. (Ed.), *Language change and functional explanations* (pp. 9–19). Berlin: Mouton de Gruyter.

Keysers, C., & Gazzola, V. (2010). Social neuroscience: Mirror neurons recorded in humans. *Current Biology*, *20*(8), R353–R354. doi:10.1016/j.cub.2010.03.013

Kim, J. (2006). Emergence: Core ideas and issues. *Synthese*, *151*(3), 547–559. doi:10.1007/s11229-006-9025-0

Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics*, *21*(1-2), 125–139.

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, *4*(6), 812–822. doi:10.1016/0959-4388(94)90128-7

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 11850–11857. doi:10.1073/pnas.97.22.11850

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews. Neuroscience*, *5*(11), 831–843. doi:10.1038/nrn1533

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, *10*(1), 110–120. doi:10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101. doi:10.1073/pnas.1532872100

LaBar, K., & Phelps, E. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, *9*, 490–493. doi:10.1111/1467-9280.00090

Lamm, C., Nusbaum, H. C., Meltzoff, A. N., & Decety, J. (2007). What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE*, *12*, e1292. .doi:10.1371/journal.pone.0001292

Landry, S. H. (2008). The role of parents in early childhood learning. In R. E. Tremblay, R.G. Barr, R. de Peters, & M. Boivin (Eds.), *Encyclopedia on Early Childhood Development* [online] (pp. 1-6). Montreal Quebec: Centre of Excellence for Early Childhood Development. http://www.child-encyclopedia.com/documents/LandryANGxp.pdf. (Accessed 25/07/2010).

Lawrence, L.L., & Fernald, Anne (1993). *When prosody and semantics conflict: infants' sensitivity to discrepancies between tone of voice and verbal content*. Poster session presented at the biennial meeting of the Society for Research in Child Development, New Orleans, LA.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–233. doi:10.1016/S1364-6613(99)01319-4

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, *22*, 1–75. doi:10.1017/S0140525X99001776

Liberman, A. M., & Mattingly, Ignatius G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36. doi:10.1016/0010-0277(85)90021-6

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*(5), 187–196. doi:10.1016/S1364-6613(00)01471-6

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289. doi:10.1146/annurev.psych.58.110405.085654

Liebermann, P. (2000). *Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought*. Cambridge, MA: Harvard University Press.

Liu, H.-M., Kuhl, Patricia K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech disrimination skills. *Developmental Science*, *6*, F1–F10. doi:10.1111/1467-7687.00275

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178. doi:10.1016/0010-0277(88)90035-2

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *10*(8), 750–756. doi:10.1016/j.cub.2010.02.045

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 756–766. doi:10.1037/0096-1523.24.3.756

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*(1-2), 233–243. doi:10.1016/S0167-6393(02)00106-1

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*(5827), 1002–1005. doi:10.1126/science.1136930

Ochsner, K., & Barrett, L. F. (2001). The neurscience of emotion. In Mayne, T., & Bonnano, G. (Eds.), *Emotion: Current Issues and Future Directions* (pp. 38–81). New York: Guilford.

Oztop, E., & Kawato, M., & Arbib, Michael. (2006). Mirror neurons and imitation: A computationally guided review. *Neural Networks*, *19*(3), 254–271. doi:10.1016/j.neunet.2006.02.002

Paavola, L. (2006). Maternal sensitive responsiveness, characteristics and relations to child early communicative and linguistic development. *ACTA UNIVERSITATIS OULUENSIS B Humaniora* 73. Doctoral dissertation, University of Oulu, Finland.

Papoušek, M. (1989). Determinants of responsiveness to infant vocal expression of emotional state. *Infant Behavior and Development*, *12*, 505–522. doi:10.1016/0163-6383(89)90030-1

Papoušek, M. (1992). Early ontogeny of vocal communication in parent-infant interactions. In H. Papoušek, Uwe Jürgens, & Mechthild Papoušek (Eds.), *Nonverbal vocal communication. Comparative and developmental approaches* (pp. 230-261). Cambridge: CUP and Paris: Edition de la Maison des Sciences de l'Homme.

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*(3), 361–379. doi:10.1353/lan.2004.0141

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews. Neuroscience*, *9*, 148–158. doi:10.1038/nrn2317

Petitto, L. A. (1987). On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition*, *27*, 1–52. doi:10.1016/0010-0277(87)90034-5

Pickering, M., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. doi:10.1016/j.tics.2006.12.002

Poizner, H., Bellugi, U., & Klima, E. S. (1990). Biological foundations of language: Clues from sign language. *Annual Review of Neuroscience*, *13*, 283–307. doi:10.1146/annurev.ne.13.030190.001435

Pourtois, G., Debatisse, D., Despland, P., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Research. Cognitive Brain Research*, *14*(1), 99–105. doi:10.1016/S0926-6410(02)00064-2

Preston, S., D., & de Waal, F. B. M. (2002). Empathy: It's ultimate and proximate bases. *The Behavioral and Brain Sciences*, *25*, 1–72.

Prinz, W. (1997). Perception and action planning. *The European Journal of Cognitive Psychology*, *9*(2), 129–154. doi:10.1080/713752551

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews. Neuroscience*, *6*(7), 576–582. doi:10.1038/nrn1706

Pulvermüller, F., Härle, M., & Hummel, F. (2000). Neurophysiological distinction of semantic verb categories. *Neuroreport*, *11*(12), 2789–2793. doi:10.1097/00001756-200008210-00036

Reilly, J. S., & McIntire, M. L. (1991). *WHERE SHOE: The acquisition of wh-questions in ASL. Papers and Reports in Child Language Development* (pp. 104–111). Stanford University, Department of Linguistics.

Reilly, J. S., & Seibert, L. (2003). Language and emotion. In Davidson, R., Scherer, K., & Goldsmith, H. (Eds.), *Handbook of affective sciences* (pp. 535–559). New York: Oxford University Press.

Rifkin, J. (2009). *The empathic civilization. The race to global consciousness in a world in crisis*. Cambridge: Polity Press.

Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194. doi:10.1016/S0166-2236(98)01260-0

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192. doi:10.1146/annurev.neuro.27.070203.144230

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge: Cambridge University Press.

Schirmer, A., & Kotz, S. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, *15*(8), 1135–1148. doi:10.1162/089892903322598102

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157–1162. doi:10.1126/science.1093535

Smith, E. E., & Kosslyn, S. M. (2007). *Cognitive psychology. Mind and brain*. New Jersey: Prentice Hall.

Stephan, A. (2006). The dual role of 'emergence' in the philosophy of mind and in cognitive science. *Synthese*, *151*(3), 485–498. doi:10.1007/s11229-006-9019-y

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*(3), 748–767. doi:10.1111/1467-8624.00313

Tamis-LeMonda, C. S., & Rodriguez, E. T. (2008). Parents' role in fostering young children's learning and language development. In R.E. Tremblay, R.G. Barr, R. de Peters, and M. Boivin (Eds.), *Encyclopedia on Early Childhood Development* [online] (pp. 1-10). Montreal Quebec: Centre of Excellence for Early Childhood Development. http://www.ccl-cca.ca/pdfs/ECLKC/encyclopedia/TamisLemondaRodriguezANGxpCSAJELanguage.pdf. Accessed 25/07/2010.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tzourio-Mazoyer, N., de Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *NeuroImage*, *15*(2), 454–461. doi:10.1006/nimg.2001.0979

van Ijzendoorn, M. H., Dijkstra, J., & Bus, A. G. (1995). Attachment, intelligence, and language: A meta-analysis. *Social Development*, *4*(2), 115–128. doi:10.1111/j.1467-9507.1995.tb00055.x

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press. doi:10.1037/11193-000

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological Bulletin*, *121*, 1–20. doi:10.1037/0033-2909.121.3.437

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expression by young infants. *Infant Behavior and Development*, *6*, 491–498. doi:10.1016/S0163-6383(83)90331-4

Werker, J. F. (1993). Developmental changes in cross-language speech perception: implications for cognitive models of speech processing. In G. Altman and R. Shillcock (Eds.), *Cognitive models of speech processing. The second Sperlonga Meeting* (pp. 57-78). Hillsdale, NJ: Erlbaum.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63. doi:10.1016/S0163-6383(84)80022-3

Zelazo, P. D., Chandler, M., & Crone, Eveline (Eds) (2009). *Developmental social cognitive neuroscience*. New York, Hove: Psychology Press.

KEY TERMS AND DEFINITIONS

Emotional Context of Learning Hypothesis: According to Harris et al. (2006), it is the emotional context of early L1 or later L2 learning that is decisive for the success of language learning. If language is learned in an emotional context, additional brain areas like the amygdala will support and thus foster linguistic representation and processing.

Emotional Reactivity Hypothesis: According to Hare and Tomasello (2005), human cognitive abilities arose through a two-step evolutionary process. First, humans learned to suppress their strong emotional reactivity towards their conspecifics for the sake of cooperation. This first step amounts to a kind of "self-domestication". In the second step, which is contingent upon the first step, higher cognitive abilities such as theory of mind, language, and problem solving, became available for selection.

Intentionality Model of Language Acquisition: According to Bloom and Tinker (2001), infants acquiring a language wish to share their thoughts about the world with their social partners. (The "aboutness" of thoughts is called "intentional".) In order to do so infants make (i) a cognitive effort to learn the language and become (ii) socially and affectively engaged with their partners.

Modality: Here, modality refers to the medium through which language is conveyed. Spoken languages use the acoustic-vocal modality; sign languages the visual-gestural modality.

Motherese: also called **infant-** or **child-directed speech (ID, CD speech):** Motherese refers to a special speech register used in interaction with infants and young children. It is characterized by exaggerated intonation and phonemic contrasts, clear segmentation clues, strong affective vocal and facial expression, among others. Motherese is thought to support language development though its efficiency is still a matter of debate.

Native Language Magnet (-Extended): According to Kuhl (1993), infants in their first year of life build up phonemic prototypes of their native language. Incoming matching sounds are attracted to these prototypes that function like perceptual magnets. Recently, Kuhl (2007) extended the earlier version of this concept to include a **gating** of this process through social aspects. Information processing in the brain is canalized through social (and emotional) factors: A socially and affectively supportive environment is deemed to be beneficial for cognitive and language development.

Paralanguage: Non-verbal, vocal aspects of language, such as affective prosody. Paralanguage is mostly processed unconsciously along with the core "verbal" aspects of language (form and meaning).

Prosody: Supra-segmental, intonational aspects of language such as temporal structure, loudness, roughness, and pitch. Prosodic contours can span parts of speech and thus help segmenting the speech stream. Prosody also exists in sign languages in the form of non-manual markers such as facial expressions and bodily movements. Prosody serves core linguistic functions (e.g., syntax) as well as affective functions.

Social Brain: Network in the brain for processing social stimuli. It comprises the processing of faces, emotion, biological motion, human goal-directed actions, and joint attention (Grossmann & Johnson, 2007). Mirror neurons are involved in the processing of goal-directed actions.

Social Referencing: In a novel and potentially frightening situation, children tend to refer back to their caregivers in order to resolve the emotional ambiguity of the situation. The "social referencing paradigm" exposes children to such emotionally charged situations and then observes their reactions.

ENDNOTES

1 Similarly, the concept of empathy is currently being revived. Rifkin (2009) even heralds an "empathic civilization", i.e., a global society whose members strive for mutual understanding and cooperation on the basis of an enlarged consciousness that includes the self and others. It is noteworthy that emotion, empathy and related topics emerge in the scientific and the societal discourse at the same time. This concurrence reflects the depth and scope of the ongoing paradigm shift in science and society.

2 Often, in the literature, the two terms "sensitivity" and "responsiveness" are used interchangeably. If a distinction is to be made, then responsiveness is related to the contingency and frequency of the mother's response toward her child whereas sensitivity is related to the qualitative appropriateness of her response in terms of the child's developmental level and situational needs (Paavola, 2006, p. 19).

3 Recently, a parallel of motherese has been argued for in the motor domain. Brand, Baldwin, & Ashburn (2002) observed that sensitive parents modify their actions when manipulating objects so as to highlight the objects' properties and affordances for the child. This sensitive enhancement has been called "infant-directed action (IDA)" or "motionese".

4 The CARE-Index as an instrument for assessing mother-child interaction was applied at 10 months of age. In addition, maternal and infant communicative acts, maternal verbal responsiveness and infant intentionality were measured. At 12 months of age, the MacArthur Communicative Development Inventories (MCDI) and the Communication and Symbolic Behaviour Scales (CSBS) were measured. Finally, at 30 months of age, the Reynell Developmental Language Scales III were measured.

5 Three of the added icons in this redrawing are standard graphics items from the word program. The icon of the brain is taken from www.schulbilder.org/gehirn-obenansicht-t4300.jpg (accessed 25/07/2010).

6 A positive language stimulus was, e.g., "Nice play" and positive paralinguistic stimulus were a smiling face and approving voice. A negative language stimulus was, e.g., "Don't touch" and negative paralinguistic stimuli were a frowning face and disapproving voice.

7 A similar discontinuous developmental pattern has been reported for the acquisition of pronouns in American Sign Language (Petitto, 1987). Sign language pronouns are identical to deictic pointing gestures, however, are subject to linguistic constraints which have to be acquired. While in the beginning, pointing gestures occur frequently, at some time in the language acquisition process they vanish completely. During this time the children work out the linguistic pronominal system. Afterwards, deictic pointing reoccurs, along with but functionally segregated from true linguistic pronouns.

8 Mukamel et al. (2010) found mirror neurons in a wider variety of brain areas than previ-

ously thought, e.g., in the hippocampus. They also found "anti-mirror neurons" that are excited during action execution but inhibited during action perception. The significance of these new findings needs to be followed up by future research.

9 Ironically, this view converges with Fodor's (1975) "language of thought" or "mentalese" hypothesis, namely the idea that human thought is structured syntactically, much like language.

10 Already very young infants have clearly differentiated expectations how animate agents and inanimate objects behave. Thus, Falck-Ytter, Gredebäck, and von Hofsten (2006) could show in an eye-tracking study that 12-month olds, but not yet 6-month-olds, anticipate the movement of a human hand towards a spatial target with their eyes pro-actively, whereas they do not show such anticipatory eye-movements if objects move to this spatial target by themselves. The authors argue that this difference is evidence for the emerging mirror neuron system during the second half of the first year of life.

APPENDIX

AFFECT IN LANGUAGE DEVELOPMENT FROM A NEUROSCIENCE PERSPECTIVE

ERP Studies on the Processing of Emotional Prosody in Infants

Language as part of the "social brain" in the sense of Grossmann and Johnson (2007) is one expressive channel for affect. In this paragraph we will be concerned with the emotional-affective tone of a verbal message which is carried by the speech prosody. Emotional prosody is characterized by its temporal structure, amplitude (loudness), roughness, and pitch (fundamental frequency) (Grossmann, Striano, & Friederici, 2005). From previous research it is known that adults distinguish reliably between neutral, happy and angry prosody, as expressed by characteristic ERP components. Thus, the difference between neutral *vs*. happy messages is captured by the higher amplitude of the P200 for happy messages, whereas the difference between neutral *vs*. happy and angry messages is captured in the N400 (as summarized in Grossmann et al., 2005). The follow-up question that could only recently be addressed since the use of ERP in infant research has become feasible is how infants process emotional prosody (Grossmann et al., 2005, 2006). Previously, the processing of emotional prosody could only be studied behaviorally. Based on these early studies, Walker-Andrews (1997) proposed a developmental sequence in which children learn to discriminate emotional expressions, first on the basis of multimodal, then prosodic, and finally facial, cues. Although the auditory system develops earlier than the visual system, prosody alone is not enough to distinguish the conveyed emotion initially. When 5-month old infants are habituated to a congruent combination of facial and vocal cues (happy/sad face plus happy/sad voice) they dishabituate if subsequently the prosody changes (happy/sad face plus sad/happy voice), thus creating an incongruity between facial and vocal information; 3-month olds can only detect the change if a sad voice changes to a happy one (Walker-Andrews & Grolnick, 1983). Grossmann et al. (2005, 2006) recently studied infants' understanding of emotional prosody with the ERP paradigm. ERP had already proven successful in the study of facial emotional expressions in adults. Schirmer and Kotz (2003) could show that a stronger negative ERP component was present when 7-month old infants looked at angry as compared to happy faces.

In a first ERP study, Grossmann et al. (2005) asked whether (i) 7-month old infants would discriminate neutral *vs*. emotional prosody (happy and angry) and (ii) among the emotional prosody, whether they would discriminate happy *vs*. angry prosody of semantically neutral German verbs. They found a stronger negative shift at 300-600 ms post stimulus following angry as opposed to happy and neutral words and a stronger positive slow wave at around 500 ms following angry and happy as opposed to neutral stimuli, however only at left temporal sites. Their results confirmed the earlier results on facial emotional processing where infants likewise responded stronger to angry as opposed to happy faces. In both cases, the processing of verbal and visual emotional information, the higher amplitude of the negative component was thought to reflect a higher allocation of attention to the negative stimuli. The ERP results confirm a more generally found "negativity bias" that has been reliably found in adults also. Taken together, these findings converge with an evolutionary viability explanation: it is much more important to detect and react to potentially threatening, negative stimuli in the environment than to positive ones. The stronger positive slow wave to both angry and happy as opposed to neutral prosody reflects a heightened sensory response to emotionally laden as opposed to neutral words in the associative auditory cortex, more specifically, in the left superior temporal sulcus. This early sensitivity to prosodic cues may underlie the finding that vocal affect, sometimes also called "paralanguage", can facilitate the recognition and learning of spoken words (see above, the studies of Friend).

In a second ERP study, Grossmann, Striano, and Friederici (2006) asked how multisensory emotional information (face and voice) is integrated in young infants. As is known from previous studies, adults readily integrate both modalities, even when they are asked explicitly to ignore one – either the facial or the vocal information. When judging faces in the presence of incongruent vocal affect, adults showed a bias towards the incongruent voice (likewise for evaluation of the voice in the presence of incongruent facial affect) (de Gelder & Vroomen, 2000). In an ERP study by Pourtois, de Gelder, Vroomen, Rossion, and Crommelick (2000), subjects first saw an angry or a sad face followed by a congruent or incongruent vocal affective stimulus which they were supposed to evaluate. The congruent condition resulted in a stronger N100 in the auditory cortex, which suggests that the facial information strengthened the auditory processing. In a second experiment, Pourtois, Debatisse, Despland, and de Gelder (2002) compared happy *vs*. fearful faces and voices, respectively. They found that congruent face-voice pairs led to a positive ERP component that had an earlier peak as compared to incongruent pairs. The site of this effect was the anterior cingulate cortex, a region that is known to play a role in error monitoring. From these results, the authors concluded that emotionally congruent information is processed faster than incongruent information. Against the backdrop of these findings in adults, Grossmann et al. (2006) conducted an ERP study with 7-month old infants on the processing of congruent/incongruent visual and vocal affective stimuli. Subjects saw pictures of a woman with a happy or angry face and concurrently heard semantically neutral words spoken with a happy or angry voice. Incongruent voices led to a more negative component as compared to congruent voices, already 350 ms after stimulus onset and with a peak at 500 ms over frontal and central electrodes on both hemispheres. Specifically, the angry voice presented with the happy face had higher negative amplitude than the happy voice presented with the angry face. Also, congruent voices led to higher positive amplitude as compared to congruent voices after 600-1000 ms over central and parietal electrodes. The authors concluded that already 7-month old infants recognize identical affect across modalities, in accordance with behavioral studies. This recognition is presumably supported by infants' memory of happy and angry faces. Since before the age of 10 months they are more likely to be exposed to happy faces, infants more strongly expect a happy voice, so that they are more surprised when they rather hear an angry voice. This reasoning could explain the above result that their response was stronger to a happy face/angry voice than to an angry face/happy voice. In this context, a study by Tzourio-Mazoyer, de Schonen, Crivello, Reutter, Aujard, and Mazoyer (2002) deserves mentioning. They found that 2-month olds activated a similar neural network as adults when viewing faces (right-lateral fusiform face area, bilateral inferior occipital and parietal areas), however, they also showed activation in one additional area, namely Broca's area. The authors speculated that this co-activation of face and language regions in young infants reflects social aspects of language in face-to-face communication and possibly serves to facilitate early language acquisition, as discussed above.

The Human Mirror Neuron System: Human Action, Emotion, Empathy, and Language

In this section, I will give an overview over studies on the adult human mirror neuron system in order to provide a relevant neuro-scientific frame for the main developmental topic of this chapter. The discovery of the mirror neuron system by Rizzolatti and his collaborators was one of the major discoveries in the neurosciences in the 1990ies. The mirror neuron system had first been found in monkeys, and later also in humans (for an overview, see Rizzolatti & Craighero, 2004). Mirror neurons are a special class of neurons that fire when a subject perceives a goal-directed action as well as when the subject plans and produces the action herself. In the monkey, this property of mirror neurons could be proven through single cell recordings (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996, among others). In humans, the evidence is mostly indirect and stems from fMRI studies (as summarized by Rizzolatti & Craighero, 2004). Mirror neurons are located in a variety of areas in the brain, in humans most notably in the left inferior frontal gyrus (IFG) that also comprises Broca's area, the classical area for language (production), in the adjacent premotor cortex and in the inferior parietal cortex. Only recently, single cell recordings have been reported in patients to be surgically treated for epilepsy (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010).[8] The primary function of mirror neurons has first been related with understanding of goal-directed transitive actions (actions directed at a goal-object) and later also with understanding of intentions (Iacoboni, Molnar-Szakacs, Gallese, Buccino, Mazziotta, & Rizzolatti, 2005). In humans, the mirror mechanism further suggests itself as an explanation of imitation (Meltzoff & Prinz, 2002; Brass & Heyes, 2005; Oztop, Kawato & Arbib, 2006). Theories of imitation presuppose some common representational format between perception and action, a "common code" (Prinz, 1997), and it is this commonality that the mirror neuron system can naturally account for. After the discovery of mirror neurons for perception and production of goal-directed, intentional behavior, their involvement in other cognitive areas such as social cognition, emotion and empathy, theory of mind, and language, has been investigated, too (Arbib, 2005; Bastiaansen, Thioux, & Keysers, 2009; Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Decety & Ickes, 2009; Fogassi & Ferrari, 2007; Gallese, Keysers, & Rizzolatti, 2004; Gallese, 2008; Rizzolatti & Arbib, 1998; Rizzolatti, Fogassi, & Gallese, 2006). The universality, immediacy, and intuitiveness with which humans understand the emotional states of others and empathize with them led to the search for dedicated neural networks supporting those (Carr et al., 2003; deVignemont & Singer, 2006; Singer, Seymour, O'Doherty, Kaube, Dolan, & Frith, 2004; Lamm, Nusbaum, Meltzoff, & Decety, 2007). One system that has been invoked for emotion and empathy is the mirror neuron system (Preston & de Waal, 2002; Decety & Jackson, 2004, 2006; Decety & Ickes, 2009; Gallese, 2001, 2008; Kaplan & Iacoboni 2006). Emotions, according to Rizzolatti et al. (2006), allow for a seemingly direct mapping between the sensory input, i.e., the observation of an emotion, and the experience of that emotion in the perceiver. The mirror neuron system thus forms the basis for empathy in the case of reading another person's affective state and for mind reading in the case of discerning the intentions of others through observation of their goal-directed behavior and emotional display. The relation between action and emotion understanding and empathy, however, is highly intricate and far from being fully understood. Pieces of a still to be uncovered mosaic are emerging, though. Kaplan and Iacoboni (2006) argued that intention understanding in the observation of human action is informed by contextual information (whether a grasping action is carried out in the context of a meal or cleaning up) as well as by specific information about the grasping type (precision *vs*. whole hand grip). Furthermore, they showed that signal changes in the relevant right posterior inferior frontal gyrus (right posterior IFG) correlated with subjective ratings in some subscales of the Interpersonal Reactivity Index (IRI), namely the fantasy scale measuring cognitive aspects of empathy and empathic concern measuring affective aspects of empathy, respectively. From these results they deduce "a central role of the human mirror neuron system in social competence." (Kaplan & Iacoboni, 2006, p. 182). In another fMRI study, Carr et al. (2003) monitored the brain activity of subjects whose task was to imitate or just observe emotional facial expressions. They found a similar network of active areas in imitation and observation consisting of the premotor face area, the dorsal sector the pars opercularis of the IFG, superior temporal sulcus (STS), the insula and the amygdala. Rather than claiming that empathy is a mirror system itself, they emphasize the link between emotion understanding and empathy with the mirror system of action understanding. By simulating the action which leads to the emotional (face) expression, the observer can empathize with that emotion. This "empathic resonance occurs via communication between action representation networks and limbic areas provided by the insula." (Carr et al. 2003, p. 5502) For the understanding of emotional facial expressions, "embodied" approaches have been put forward. Either mimicry or internal simulation is invoked as a mechanism for understanding others' emotions. A direct account in terms of facial mimicry – which may, however, remain subliminal – is proposed by Niedenthal (2007). She reports behavioral studies measuring activation of facial muscles corresponding to emotions in subjects looking at pictures that evoke various emotions: anger was related to a subliminal activation of the corrugator supercilii ("frowning"), disgust to the levator labii ("grimacing"), and joy to the zygomatic major ("smiling"). Other studies showed that reading words that are related to positive and negative emotions, e.g., "to smile" or "to frown", and that involve typical facial expressions, leads to motor resonance *via* the subliminal activation of the corresponding muscles (zygomatic major and corrugator supercilii) (Foroni & Semin, 2009). This effect is stronger for verbs than for adjectives that are more abstract and less directly related to facial expressions ("funny", "annoying"). The effect vanishes if motor resonance is inhibited as when subjects have to hold a pencil between their lips so that the muscle cannot engage in motor resonance when the emotional word is being processed. Not only the conscious perception of emotion words but also their unconscious processing induced by subliminal presentation evokes motor resonance. Motor resonance was shown to further influence evaluative judgments of cartoons. Thus, subjects whose "smile" muscles are not disabled, reading the word "smile" on a computer screen subsequently rate cartoons funnier as compared to subjects whose "smile" muscles are inhibited. Foroni and Semin (2009) as well as Niedenthal (2007) take these results as evidence for the "embodiment of language" in general, and for the "indexical hypothesis of language comprehension" of Glenberg and Robertson (1999, 2000), in particular. Somatic responses involved in the processing of emotion words lead to simulation of the emotional experience in the perceiver and thus contribute to the understanding of the meaning of those words. How exactly the concepts of facial mimicry and motor resonance are related to the mirror neuron mechanism in the brain is not fully clear, yet. A more indirect account in terms of internal embodied simulation (and not just overt facial mimicry) that makes the hidden inner states of the observee accessible to the observer is proposed by Bastiaansen et al. (2009). They call this link the "Rosetta stone" (2009, p. 2397) that helps translate observable (facial) actions of others into their hidden internal states, i.e., their emotions and intentions. This link could exist between the premotor cortex and IFG. Since premotor areas do not activate motor actions (such as facial expressions), this more indirect account seems less embodied and somewhat more cognitive than the more direct account of Niedenthal, Foroni and Semin. However, common to both approaches is the claim that it is the modality-specific, perceptual systems of the brain that support the simulation of the affect and that simulation is not so much a matter of the muscles and viscera (Niedenthal, 2007, p. 1005).

Chakrabarti, Bullmore, and Baron-Cohen (2006) showed that a cluster of mirror-neuron areas, namely left dorsal inferior frontal gyrus and premotor cortex, correlated positively with the "Empathy Quotient" (EQ), a measure of trait empathy, across a variety of emotional conditions (happy, sad, angry and surprised facial expressions). Thus, the mirror neuron system mediates between the perception and recognition of actions and emotions. Also de Vignemont and Singer (2006) stress the link between empathy and action understanding. They propose that "empathy provides a more precise and direct estimate of other people's future actions because shared emotional networks also directly elicit the activation of associated relevant motivational and action systems." (p. 439). This is an epistemological function. A social function of empathy is that it facilitates social communication and social coherence. The link between the presumed mirror neuron system for emotion and communication points to a possible connection between the mirror neuron system and language. Already long before the discovery of mirror neurons, it had been proposed in the "motor theory of speech perception" that the sounds of language are perceived in terms of their motor gestures (Liberman & Mattingly, 1985; Liberman & Whalen, 2000). After the discovery of mirror neurons, a mirror mechanism was proposed for the relation between language perception and production, too (Arbib 2005; Gallese 2006; Oztop, Kawato, and Arbib 2006; Rizzolatti and Arbib 1998). In a similar vein, simulation of language production during comprehension has been invoked as a mechanism for explaining the ease of mutual understanding in communication (Pickering & Garrod, 2007). While these studies focused on the inherent mirror qualities of the language system, the relation between emotion and language in the scope of a mirror mechanism became a research topic subsequently. Not only direct observation of (visual) emotional displays, as discussed above, informs us about the affective state of others, but also properties of language. Here, language is conceived of in terms of human social cognition and not so much as an autonomous mental faculty of the human mind (in the sense of Chomsky, 2000, 2007, among many others). In such an "embodied view" on language, the support of human social communication and action by the language circuits in the brain are in focus (Gallese, 2008; Niedenthal, 2007). Crucial in this respect is the re-assessment of the role of Broca's are (BA 44), which is one part of the left ventral premotor cortex. Previously, it had been reserved for language (production) functions only; however, nowadays it is considered to play a major role in action understanding and production also. Broca's area generally supports the construal of complex hierarchical representations that may serve the planning and production of action sequences, musical themes, and sentences alike. The notion of "syntax" thus becomes generalized and comprises various domains of cognition, not just language.[9] Mirror neurons provide a mechanism that establishes a direct relation between the sender and receiver of a communicative message, be it an observed action or communication. They solve two problems at the same time: the "parity" and the "direct comprehension" problem (Rizzolatti et al., 2006). Parity means that the meaning is the same for the sender and the receiver and direct comprehension means that there is a direct, hard-wired mapping of the percept (visual action scene, emotional expression) onto its meaning, "without any cognitive mediation" (Rizzolatti & Craighero, 2004, p. 183). Turning to the relevance of the mirror neuron account for language, Gallese (2008) proposes an "embodied" view that grounds the processing of language in the human action system. He points out how such an embodied language is instantiated on various levels: the vehicle level of motor articulation, the content level of meaning, and even the syntactic level of recursive phrase structure. In his "neural exploitation hypothesis" he claims that neural circuits that originally evolved for the perception and control of motor actions were later exploited by the newly emerging language faculty. Thus, language (and thought) inherited crucial properties, among them "mirror" properties, from these evolutionarily older action systems. Much like stringing together individual actions to form an "action sentence", we are now also stringing words together to form sentences. A crucial "relay" mechanism might be provided by the "auditory mirror neuron system". Shortly after the discovery of visual mirror neurons another population of mirror neurons was found (in the monkey brain) that respond to the production of an action and also to the perception of the specific sounds that accompanied actions and instantiate action effects, e.g., cracking a peanut, tearing a sheet of paper, or opening a can (Rizzolatti & Craighero, 2004; Gallese, 2008; Gazzola et al., 2006; Aziz-Zadeh et al., 2010). Some of these actions and corresponding sounds involve the hand and some of them the mouth. The left temporo-parietal premotor circuit is organized in a somatotopic way, i.e., hand actions/sounds activate dorsal premotor hand areas (BA6) and mouth actions/sounds activate left ventral premotor mouth areas (BA 44) (Gazzola et al., 2006). This somatotopic organization is also supported by the effector-specific "word nets" that Pulvermüller and colleagues found for action words related to those effectors. They showed in a number of studies that during the processing of action words that imply certain body parts such as "lick", "pick", and "kick" (pre-)motor areas of the brain, namely for the mouth, hands, and legs, are co-activated along with the classical language areas. Likewise, during the processing of visual words such as "see", areas in the occipital lobe are co-activated (Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, 2005, among many others). Since the auditory mirror neuron system is left-lateralized, as expected from its relation with language, it may also be connected to the left-hemispheric multi-modal mirror system that comprises the visual modality, thus forming a supra-modal interface between human action and language in perception and production (Gazzola et al., 2006, p. 1827).

Emotional resonance, as suggested by the "embodied" or simulation account of emotion understanding, has been related to observational learning (Niedenthal, 2007). If the emotional expression of the observed other person is mirrored in the observer, the corresponding emotional experience may be relived and thus an empathic understanding may be reached. Not only through observational learning, but also through instructed learning, i.e., through language, reexperience of an emotion may be triggered. As Niedenthal (2007, p. 1004) points out, a child being told not to put her fingers into an electrical outlet, otherwise she will experience a painful electric shock, must be able to reexperience the linguistically conveyed emotion (here, pain).

Few neuro-imaging studies have so far investigated the production of emotional words. Cato et al. (2004) showed in an fMRI study that (silently) generating emotional words with positive and negative meaning activates dedicated areas in the rostral frontal and retrosplenial/posterior cingulate cortex as compared to emotionally neutral words. The rostral, frontal area presumably supports "the generation of words with emotional connotation", whereas the role of the retrosplenial/posterior cingulate cortex is "the evaluation of an external stimulus with emotional salience" (Cato et al., 2004, p. 173). More common are studies that investigate the brain areas active in the production of syllables with specific affective prosody. These studies tap into the relation between affective aspects of language and empathy.

As the mirror neuron system is presumably in the service of understanding others, several neuro-imaging studies on the perception (and production) of human action and action-related sounds also assessed individual empathy levels of the subjects. The hypotheses (i) that there should be a relation between the activation of the mirror neuron system and empathy, as measured in questionnaires in general, and, more specifically, (ii) that higher empathy ratings should go along with higher levels of activation in the mirror neuron areas were confirmed in several studies (Gazzola et al., 2006; Aziz-Zadeh et al., 2010; Kaplan et al., 2006). The property of language most intimately related to the expression of affect is prosody. As prosody is related to language and to emotion, for both of which mirror neuron systems had been found, it was conjectured that there may exist a mirror system for prosody, too. A candidate system would be the auditory mirror system as discussed above. In a combined behavioral and fMRI study, Aziz-Zadeh et al. (2010) assessed subjects' brain activity while they judged the emotional quality of audio-clips with happy, sad, question, and neutral intonation. Also, subjects had to utter non-sense syllables ('dadadadada') with emotional prosody in the scanner. In the behavioral part, they tested again subjects' own production of affective prosody (that was subsequently rated for their prosodic quality) and had subjects rate another set of audio-clips with affective prosody. In order to assess their empathy, subjects filled in a two questionnaires. The fRMI study showed that common areas were active during perception and production of emotional (as well as linguistic) prosody, namely areas in the left premotor cortex (left inferior frontal gyrus (IFG), left dorsal premotor cortex). This is evidence for a prosodic mirror system. IFG and premotor areas have links to auditory areas. Through this dorsal pathway the interface between perceived and articulated language might be constituted. Moreover, there were correlations between the amount of activation of brain areas for emotional prosody and empathy scores as well as between prosodic perceptual abilities and empathy scores. Also prosody production inside and outside the scanner were found to be correlated. From these findings Aziz-Zadeh et al. concluded: "This data support the notion that components of empathy to emotional stimuli may rely on simulation processes carried out, in part, by motor-related areas (…). Thus, in order to understand someone else's prosodic intonation, we may simulate how we would produce the given intonation ourselves, which in turn may be a component of the process involved in creating empathic feeling for that individual." (2010, p. 6)

Motor resonance might equally contribute to and facilitate lexical learning in young children. Contexts in which young children learn words are those where they interact with others and with objects that have certain characteristics, internal, and external. These properties are revealed during the interaction, e.g., emotions, motives, and goals of the human play partners on the one hand, and physical properties, affordances, and action effects of the objects, on the other hand.[10] The tight connection between emotions and language is witnessed in language acquisition where initially, in the first year of life, the speech melody that carries the affective prosody, is the main message to the child (Papoušek, 1992). Young children are more susceptible to this information in the speech stream than to the lexical information, initially. When they acquire language, however, they have to learn to integrate the emotional information and the lexical information – a not at all trivial task (see section on "Parents' sensitivity to informational and affective content in their infants' vocalizations" above).