Language acquisition and ERP approaches: prospects and challenges

David Poeppel and Akira Omaki

1 Interdisciplinary preliminaries

Developmental cognitive neuroscience is a vibrant and growing domain of research, and various aspects of human psychological experience are being investigated using techniques that bridge developmental psychology, developmental biology, the cognitive sciences, and non-invasive neurobiological techniques. Ranging from the growth of perceptual expertise to ontogenetic change in decision making, practically every part of human perception and cognition is being evaluated – and reevaluated – from a developmental perspective enriched by neurobiological methodologies (Nelson and Luciana 2001; Johnson 2005).

Rapidly expanding scientific fields offer grounds for optimism but also cause for healthy skepticism, and before we go on to discuss the important progress made in current work on the developmental cognitive neuroscience of language, we feel compelled to remind ourselves as well as the reader that the intellectual challenges remain daunting. For example, the general excitement and implicit promise of the new integrative approaches notwithstanding, it is worth considering some basic assumptions that, in our view, tend to remain unexpressed in the field. In particular, we take it to be a reasonable presupposition that a developmental cognitive neuroscience study will enrich our understanding of *either* how development works, *or* how cognition works, *or* how the brain works. This desideratum may sound innocent enough, but if we subject numerous cognitive neuroscience studies to this straightforward standard, the analysis can be sobering: a surprisingly large number of studies (in any branch of the cognitive neurosciences) do not, at least in any obvious way, contribute either to a deeper understanding of human cognition or human brain function. Rather, much work is at best correlative (which is, of course, not without value, but presumably not the ultimate goal) and occasionally even *sui generis*, not linking to either the cognitive or the brain sciences.

It goes without saying that the problems under investigation are quite difficult indeed, and it constitutes a formidable challenge to generate meaningful accounts that link a deeper understanding of the human language faculty with biological data. In part this difficulty stems from the strikingly different ontological commitments made by the different disciplines (or, more colloquially, the 'parts lists'). As discussed by Poeppel and Embick (2005), the fields of inquiry focused on language operate over such 'elementary representations and computations' as distinctive features, syllables, morphemes, noun phrases, semantic composition, or syntactic displacement operations. The neurobiological sciences, on the other hand, build on primitives that include synapse, neuron, cortical column, oscillation, and so on. In trying to build substantive connections between these differing domains of inquiry, notably absent are the linking hypotheses that would allow us to state in explicit terms how these sets of putative primitives are related. The challenges associated with this aspect of interdisciplinary research were called the "granularity mismatch problem" and the "ontological incommensurability problem" to highlight how complex it is to relate biological and cognitive levels of analysis in a manner that is beyond merely correlative (Poeppel and Embick 2005). For example, highlighting in an imaging study that a brain area is activated for syntactic processing, while interesting, remains remarkably underspecified and entirely uninformative with respect to mechanism. 'Syntactic processing' and 'lexical access' and 'lexical semantics' are not elementary and simplex cognitive operations; similarly, brain areas on the scale of a centimeter (say Brodmann's area 45) are not elementary and simplex biological structures. Overall, it must therefore be our (interdisciplinary) goal to decompose the cognitive science aspect and the neurobiological aspect in such a way as to permit fruitful and workable linking hypotheses.

While we are skeptics in stance, we remain relentlessly optimistic about the promise of discovering deep principles of human development and cognition, and we turn now to the prospects of this important area of growth in psychological research. Can cognitive neuroscience engage some of the hard problems, say of the type, "What are the neuronal mechanisms that form the basis for the representation of and computation with linguistic primitives?" In our view, the area of language development – while not immune to the dangers inherent in interdisciplinary research – is well poised to make serious and satisfying progress. We hold this to be true because (i) the questions associated with language acquisition and development are well characterized theoretically and connect to an extensive behavioral and computational literature. (ii) There are few issues more pressing and riveting (both for basic and clinical research) than discovering how human development works and constrains human perceptual, cognitive, and affective experience. Therefore, a lot of exciting new research is enriching our knowledge of developmental mechanisms in ways that ERP, for example, can build on effectively. (iii) New techniques promise to yield unexpected insights into brain function throughout the lifespan; hemodynam-

ic imaging and electrophysiology are increasingly adapted to record from developing brains and will provide new vistas onto ontogenetic change in childhood.

Looking at the field of language acquisition from a historical perspective, there have been a series of influential trends, and a useful summary is provided by Mehler and colleagues (this volume). Three trends that have been particularly consequential are (i) the acquisition of language-specific rules, an area of inquiry particularly prominent in the context of generative grammar, (ii) the acquisition of the phonetic inventory and the problem of word learning, areas of research at the interface of linguistics and cognitive psychology, and (iii) statistical learning approaches, research more squarely driven by experimental psychological concerns and methodologies. While these thrust areas have generated a tremendous amount of data on language development, they have remained largely insulated from neurobiological concerns. Despite some notable exceptions, the bulk of language acquisition research has been concerned with cognitive science, in the broadest sense – but not cognitive neuroscience. The chapters presented in the current volume represent an important new trend, namely bridging theoretically motivated questions in language development with new biologically based techniques that yield innovative measures of knowledge and performance.

Here we summarize and highlight some of the major observations outlined in the chapters (Section 2), discuss the methodological and conceptual challenges (Section 3), and make some specific recommendations to further strengthen the work on the developmental cognitive neuroscience of language (Section 4).

2   Highlights

The behavioral and electrophysiological studies reported in this volume contribute critical data to our understanding of the development of psycholinguistic processes as well as the neural activity that underlies these processes. In this section, we highlight some of the unique contributions reported in each chapter. This brief review forms the basis of the subsequent sections regarding the prospects and challenges for cognitive neuroscience research on language development.

Four areas of research receive special emphasis in the chapters, in part reflecting domains of acquisition in which ERP research can be argued to yield promising insights. (1) Acquisition of the phonetic inventory (Conboy chapter). (2) Word segmentation and recognition (Nazzi, Kooijman, Thierry, Sheehan, Conboy chapters). (3) Lexical and conceptual semantics in single word processing (Friedrich, Sheehan chapters). (4) Structural learning and processing (Conboy, Mehler, Friederici chapters). Table 1 summarizes the essential attributes of the chapters that contribute child ERP data and can serve as the basis for a 'conceptual meta-analysis.'

Table 1. Areas of child ERP research thematized: Sounds to sentences, 3 months to 48

Domain of Acquisition	Chapter	Tasks	Languages tested	Age (months)	Major ERP responses
Phonetic inventory	Conboy	– Passive listening (CV oddball paradigm)	English, Spanish	7	P150–250 or N250–550 (native & nonnative contrasts)
				11	N250–550 (native contrasts)
				20	N250–550 (native & nonnative contrasts)
Word segmentation/ recognition	Kooijman	– Passive listening (sentences containing familiarized word)	Dutch	7	Early frontal positivity (~ 350 ms) Left-lateralized negativity (~ 480 ms)
				10	Left-lateralized negativity (~ 350 ms)
	Thierry	– Passive listening (familiar and unfamiliar words in isolation)*	English	9/10/11	N2/N4 modulation (familiar-unfamiliar)
				12	No N2/N4 modulation
			Welsh	9/10/11/12	No N2/N4 modulation
			English-Welsh bilingual	11	N2/N4 modulation in English and Welsh
	Conboy; Sheehan	– Passive listening (known and unknown words in isolation)*	English-Spanish bilingual	19–22	N200–400 (known-unknown words)
		– Passive listening to (known and unknown words in isolation)*	English	13–17, 20	N200 and N350 (known-unknown words)

Domain of Acquisition	Chapter	Tasks	Languages tested	Age (months)	Major ERP responses
	Sheehan	– Passive listening (words, infant directed speech, IDS, vs. adult directed speech, ADS)	English	6	Larger N600–800 to familiar words in IDS
				13	Larger N600–800 to all words in IDS Larger N200–400 to familiar words in IDS
		– Passive listening (familiar, unfamiliar, backwards words)	English	3–4	Positivity to all types (175–550 ms)
				6–8	N200–500 (familiar-backwards)
				9–11	N200–500 (familiar & unfamiliar-backw.)
	Friedrich	– Picture-word matching paradigm	German	12	Early negativity (congruous/incong. words)
Lexical semantics				14, 19	Early negativity N400 (congruous/incong.)
	Sheehan	– Picture-word matching paradigm	English	13/20/36	N400 (congruous/incong. words)
		– Word-pic. matching – Gesture-pic. matching	English	18	N400 word-pic & gesture-pic
				26	N400 word-pic only
Sentence structure	Conboy	– Passive listening to sentences	English	30/36/48	N400 to semantic anomaly P600 to morphosyntactic anomaly
	Friederici	– Passive listening to sentences	German	24	P600 to phrase structure violation
				32.5	P600 and ELAN

* (Un)familiar and (un)known are used in specific differing ways in these studies. Known is used in the context of verifiable parental report; in contrast, familiar is used when the lexical item in question is a word typical of children's vocabulary but not necessarily attested.

2.1   The inventory of speech sounds

A highly productive area of language acquisition research deals with speech sound perception and acquisition. Numerous behavioral studies have established that infants in their first year transition from being 'universal listeners' – with the ability to distinguish all phonetic contrasts in natural languages – to language-specific listeners (Kuhl 2004). Infants' discrimination ability declines (differentially for vowels and consonants) throughout the first year and adjusts to the phonetic inventory of the target language (Jusczyk 1997; Kuhl 2004; Werker and Tees 1999).

ERP research reviewed by Conboy and colleagues (this volume) builds on this rich source of behavioral data and sheds light on the associated neural activity. Two striking aspects of their findings are that (i) infants at 11 months and 20 months still demonstrate neural sensitivity to non-native consonant contrasts (unlike what previous behavioral research has suggested), and (ii) a subgroup of infants show different electrophysiological responses at certain ages, and the developmental trajectory of these responses in the first year of life can be used as a predictor of later language development. In particular, ERP responses such as an early (~150–250 ms) positivity, the MMN, as well as extended negativities (~250–550 ms) appear to be robustly measurable during the first year of life and therefore can serve as non-invasive metrics of perceptual performance. More generally, such data indicate that neurophysiological responses may provide more sensitive measures that can uncover cognitive processes that are not observable in the context of behavioral paradigms. Consequently, subtle neurocognitive processes underlying the acquisition of the speech sounds inventory can be carefully characterized using a mixture of new neurophysiological and established behavioral approaches.

2.2   Segmenting the signal and finding words

A large body of work has been devoted to the basic question of the recognition of known words, and specifically of finding words in continuous speech, an essential step for constructing a lexicon for the target language (Jusczyk 1997). This task is necessarily quite challenging, since words are produced in isolation only occasionally (Woodward and Aslin 1990), and furthermore because it requires the learners to be able to access the phonetic details necessary to retrieve the associated semantic memory.

Nazzi and colleagues (this volume) investigate whether children learning different languages use language-specific word segmentation strategies. They hypothesize that learners' initial segmentation algorithm is based on 'primary', rhythm-based attributes characteristic of the target language that provide cues for segmentation. They predict, therefore, that L1 learners of French, unlike those of

English, will principally use syllabic-based cues for word segmentation (versus stress-based cues typical of English rhythmicity). Their behavioral results, based on the head turn preference procedure, indicate that French 12-month-olds have a syllabic segmentation procedure, but 16-month-olds do not. However, the syllabic segmentation procedure may still be present in 16-month-olds, although masked by the use of other (now salient) segmentation cues (e.g., transitional probabilities, phonotactic information) available at this later developmental stage. Nazzi and colleagues argue, rightly, we think, that the sensitivity and temporal resolution of ERP might permit investigators to observe the early syllabic segmentation as well as the intermediate representations still available. The concurrent use of behavioral and ERP data might therefore illuminate the specific hypothesis that representations of a certain type (say, the syllable) are operative during processing at different ontogenetic stages, thereby providing a richer, more psycholinguistically specified, account of the U-shaped developmental function demonstrated by purely behavioral metrics.

Kooijman and colleagues (this volume) specifically address the problem of word segmentation in continuous speech (Jusczyk and Aslin 1995) as well as the comparability of behavioral studies and ERP studies addressing this question. Their ERP experiment adopts designs similar to previous behavioral studies. This strategy strikes us as eminently reasonable, insofar as it allows the ERP data to be interpreted in the context of theoretically motivated behavioral data. Indeed, these authors provide a thoughtful commentary on the methodological challenges associated with infant and child ERP studies (see Section 3.1 below for discussion). They show that in Dutch infants, the basis for word segmentation abilities already exists in 7-month and 10-month-olds (though they give slightly different responses). In contrast, the previous behavioral studies suggested that word segmentation abilities develop around 9 months (Houston, Jusczyk, Kuijpers, Coolen and Cutler 2000). These investigators also conducted a head turn preference experiment using a design maximally similar to that of their ERP experiment, and they still found that 7-month-olds do not demonstrate word segmentation abilities. This set of studies suggests, again, that ERP measures may be more sensitive to emerging cognitive processes in infants.

A series of studies by Thierry and Vihman (this volume), Sheehan and Mills (this volume), as well as Conboy and colleagues (this volume) examined word recognition and its underlying neural mechanisms in monolingual and bilingual children. Using the N200–400 response components elicited by the presentation of words, Sheehan and Mills tested a variety of factors that affect word recognition abilities in one- and two-year-olds (age, training, vocabulary size, ability to use phonetic details, and infant vs. adult directed speech). Bilingual studies conducted by these authors demonstrate that language experience significantly affects the

process of vocabulary development. For example, the influence of social environment on language input plays a crucial role in vocabulary development (the case of monolingual Welsh-speaking children and English-Welsh bilingual children reported by Thierry and Vihman), and language dominance is characterized by early emergence of (adult-like) focalization of EEG responses to known and unknown words (Sheehan and Mills; Conboy and colleagues).

An important common feature across these studies is the consistent identification of an event-related negativity ranging approximately from 200–500 ms post onset of the critical item. This negative ERP is likely to be the precursor of the N400, although a thorough cross-sectional study on N400 morphology and timing has, to our knowledge, not been conducted. Therefore, we must remain open to the possibility that the negativity seen in infant and child parietal ERPs is not entirely congruent with the adult N400. This response directly sheds light on the development of neural mechanisms that support lexical access, which could not be investigated by behavioral measures alone. In other words, this offers an additional dependent measure that can be used to assess with great sensitivity how the development of lexical access and lexical representation unfolds in the developing child.

2.3   The meaning of words

Studies reviewed in Friedrich (this volume) as well as Sheehan and Mills (this volume) investigate the development of lexical semantics using a picture-word matching paradigm. In these types of experiments, the presentation of a picture is followed by a congruous or incongruous word. Friedrich reports two ERP components: an early lateral frontal negativity that apparently reflects facilitation of word form processing, and an N400 that reflects semantic integration processes. Friedrich finds that both of these components are present in German-speaking 14-month and 19-month-olds, whereas only the early negativity is elicited in 12-month-olds. Sheehan and Mills used a similar experimental procedure with English-speaking infants, and found that even 13-month-olds showed an N400-like effect, although it differed slightly in latency and duration from that reported for older children (20-month and 36-month-olds). It is thus not clear exactly when the N400 starts to emerge in infants' lexical processing, but these studies undoubtedly show that an N400-like response is robustly elicited sometime during the second year of life, a period which is characterized by rapid vocabulary growth. An important common theme across these studies, one that resonates with much current research in sentence processing, is the concept of *prediction* in language processing. What the studies show convincingly is that the presentation of prior information systematically constrains which lexical representations are suitable candidates for a given context. The idea that a speaker/hearer entertains a detailed and grammatically so-

phisticated current model that is used at every processing step (*analysis-by-synthesis*) is gaining acceptance in areas of research ranging from speech perception (Poeppel, Idsardi and van Wassenhove in press) to sentence processing (Phillips and Wagers in press) and even to visual object recognition (Yuille and Kersten 2006).

2.4 Finding and processing structure in sentences

Past decades of linguistic research have established that natural language is characterized by a computational system that manipulates abstract linguistic representations (Chomsky 1957). Although this line of inquiry has proven to be productive, it naturally raises the following two questions. First, the characterization of linguistic knowledge forces a learnability question, as such abstract rules and representations are not directly observable in the input (Pinker 1984). This issue has led to proposals of innate constraints on the hypothesis space (Crain and Thornton 1998) as well as powerful (but sufficiently constrained) learning mechanisms that allow children to induce abstract generalizations (Elman 1993; Newport and Aslin 2000). While children are learning their native language(s), they must be able to analyze and internally represent the input to enable comprehension. This raises the second question: to what extent are children capable of computing such structural representations during development?

Mehler and colleagues (this volume) present a critical survey of the recent literature on artificial language learning, a field that has generated provocative debate about the nature of children's learning mechanisms. Although the discussions tend to focus on the comparison of learning performance by a (constrained) statistical learner (Newport and Aslin 2000; Saffran, Aslin and Newport 1996) and a rule-extraction/generalization learner (Marcus, Vijayan, Rao and Vishton 1999), Mehler and colleagues argue for the importance of an alternative perspective, namely, a 'perceptual' bootstrapping mechanism that is attuned to 'perceptual primitives': they propose an operation sensitive to identity relations and an operation that detects and uses the information at the edge of a given representational unit; both of these operations are argued to be used extensively by the language learner. These authors suggest that the interaction of statistical computations and those specific perceptual primitives can feed into mechanisms dealing with more abstract structures and lead eventually to learning the relevant abstract rules and representations.

Conboy and colleagues (this volume) as well as Friederici and Oberecker (this volume) explicitly investigate, using ERP, the sentence processing mechanisms testable in 2 to 4-year-old children. Conboy and colleagues show that English-speaking children as young as 30-months show adult-like N400 responses to semantically anomalous sentences and P600 responses to morphosyntactic anoma-

lies. Friederici and Oberecker, in turn, use a version of phrase structure violation sentences that are widely studied in adult German ERP research (for a review, see Friederici 2002; Friederici and Weissenborn 2007). They observe that 32.5-month-olds show adult-like biphasic ELAN and P600 responses, whereas only the P600 is elicited in 24-month-olds. These findings indicate that children's parsers (and, by extension, grammars) are qualitatively the same as adults'. Indeed, these authors argue that the ERP data recorded in the context of sentence processing experiments in children are most consistent with a 'continuity' perspective on language development. If such arguments are on the right track, it would suggest that these neurobiological data can be used to adjudicate between theoretical alternatives about child development.

3   Prospects and challenges: The state-of-the -art

3.1   The technological challenge

The main feature of the present volume concerns the technical advances seen in the use of EEG with language learners. At first glance, ERP offers the best of all possible worlds: its temporal resolution (msec) is appropriate to the phenomena under investigation. Whether in the study of speech perception (Conboy, this volume), lexical processing (Nazzi; Kooijman; Thierry and Vihman; Sheehan and Mills; Friedrich; all this volume), or sentence processing (Conboy; Friederici and Oberecker; this volume), the operations that underlie perceptual and linguistic computation are extremely fast, typically transient, and follow each other at rapid intervals (see, e.g. Friederici 2002 for review). ERP is well suited to capture the relevant processes. Moreover, the task requirements are suitable, permitting passive presentation, and requiring no overt tasks of infants and toddlers. We review below some methodological concerns inherent in the use of EEG measures with children, and discuss how one might make use of behavioral and EEG measures to contribute critical data.

3.1.1   *Fine time course measure and ERP components*
Multiple cognitive processes underlying language comprehension occur within a few hundred milliseconds of stimulus presentation, and a dependent measure that records responses within an appropriate time frame may capture some of the multi-level processes, such as speech perception, word recognition, or integration of the lexical item into a syntactic structure. For this reason, using a behavioral response alone it can be difficult to pinpoint which of these underlying subroutines are responsible for observed differences between conditions. Furthermore, since

standard reaction time measures are mediated by other motor responses such as eye movements or button pressing, the temporal precision of the underlying cognitive processes is rather poor. Recording of EEG, however, allows us to inspect directly the neural activity that underlies the target cognitive processes, and hence provides us with extremely good temporal precision. Moreover, different ERP components are often associated with distinct cognitive processes, and this can facilitate distinguishing which of the multiple processes that occur in parallel were affected by the experimental manipulation. The utility of a precise time course measure was highlighted by Nazzi and colleagues (this volume). Their hypothesis regarding whether 16-month-olds still have a syllabic-segmentation strategy crucially relies on evidence from the intermediate representation being built before the whole word is recognized, and it seems very reasonable to use ERP to investigate this fast, cascaded sub-processes of word recognition.

The polarity and topographic information derived from ERP components turns out to be very informative in child ERP research as well. Conboy and colleagues (this volume) report that differential ERP responses (either P150–250 or N250–550) to non-native phonetic contrasts in the first-year of life can be a predictor of later language development. Since the discrimination behavior is presumably attested in both P-responders and N-responders, the polarity information provided by ERP has the potential of disclosing properties of child language that were not observable in behavioral measures. With respect to the scalp distribution of ERP components, Sheehan and Mills (this volume) report a word learning experiment with 20-month olds, in which the infants were exposed to novel word-object associations. They found that the bilateral N200-N500 component that is elicited in response to familiar words becomes left-lateralized after further training, although this lateralization effect was observable only in children with relatively larger vocabulary size. This suggests that lexical processing becomes more specialized and lateralized towards the left hemisphere as children become much more efficient word learners. Cumulatively, these studies demonstrate how informative ERP responses can be with respect to the underlying cognitive processes that are not directly observable in behavioral measures such as looking time, head-turn preference, or high-amplitude sucking.

3.1.2 *Higher sensitivity of ERP responses to underlying cognitive processes*
Some researchers show that ERPs can reveal evidence of stimulus discrimination even when behavioral measures indicate otherwise. For example, ERP data indicate the presence of non-native phonetic contrast discrimination at a later age than previously thought (Conboy and colleagues) or an early emergence of word segmentation and recognition (Kooijman and colleagues; Thierry and Vihman), as reviewed in Section 2.

The precise reasons for such behavioral vs. ERP differences are not clear, but one possible factor is the temporal resolution discussed above: Since ERPs can directly tap into the early stages of processing, they can reveal processes that are later concealed by other cognitive operations (cf. Nazzi, this volume). Another possible factor is the difference in task demands and cognitive load: Most of the behavioral measures are attention-dependent, namely, children must pay sufficient attention to the linguistic or visual stimuli to generate looking time differences. However, ERP recording does not require explicit attention or overt responses. This may reduce the cognitive load on children, and therefore ERP approaches may allow us to observe cognitive processes in younger children who may not yet have sufficient cognitive resources to handle the task demands.

3.1.3   *Applicability across different age groups*
Since ERP responses are automatically elicited upon exposure to linguistic stimuli, processing at various levels can be tested across different age groups. This has allowed developmental researchers to investigate the developmental time course of language comprehension mechanisms for phonetic, lexical and sentence processing (for a review, see Friederici 2005). This line of research has shown similarities and differences in children and adults' language comprehension mechanisms, which raise the well known continuity question: Does a child's language comprehension mechanism qualitatively differ from that of an adult? Even though a continuity question in language development has been addressed in investigations of linguistic competence (e.g., Crain and Thornton 1998; Pinker 1984), it has not been addressed nearly as much with respect the performance mechanisms, mainly because it had to await the advent of on-line measures that can be adapted to children. Thus, ERP can now be used to examine the developmental trajectory of, for example, the sentence processing mechanism (as reviewed in Section 2.4).

3.1.4   *Caveats*
Despite these appealing advantages of ERP measures, the interpretation of child ERP data requires special caution. Both Kooijman, Johnson and Cutler (this volume) and Sheehan and Mills (this volume) provide thoughtful commentary on the methodological challenges. We add some considerations here. First, we are still in need of much more basic understanding of the relation between brain development and ERP components. For example, in adult ERP research, the N1-P2 complex has been shown to be a robust, automatic response to auditory stimuli, but this complex is immature and its development lasts up to mid-puberty (Pasman, Rotteveel, Maassen and Visco 1999; Pang and Taylor 2000). As Kooijman *et al* (this volume) note, this suggests the possibility that ERP findings from children and adults may not be so directly comparable, although the fact that some ERP com-

ponents (e.g., mismatch negativity, N400, P600) have very similar properties in children and adults in terms of scalp distribution and latency intimates that at least some adult-like ERP components do exist early on in development.

Second, the fact that ERP experiments with children do not require attention to stimuli or overt responses does not mean that it is easier to acquire clean data – in fact, extremely careful artifact rejection procedures are necessary (see Sheehan and Mills, this volume, for valuable discussion). For example, researchers not only provide positive reinforcement for sitting still during periodic breaks, but also monitor the child during the experiment and mark trials on which the child did not pay attention, such that trial-by-trial basis artifact rejection techniques can be used. Furthermore, they adjust artifact rejection thresholds for each child by visually inspecting each trial, while also using computer programs to adjust thresholds based on the presence of blinking, etc. Thierry and Vihman (this volume) use a manual stimulus delivery procedure coupled with online infant monitoring to reduce the number of artifacts recorded. On balance, many conservative artifact rejection algorithms are necessary for child ERP data, and yet clean data are not guaranteed even with careful data reduction procedures.

Finally, despite the important advantages of ERP measures, behavioral studies will certainly remain critical as experimental techniques for language development research. Experiments using the head turn preference procedure (Kemler-Nelson, Jusczyk, Mandel, Myers and Turk 1995), preferential looking paradigms (Hirsh-Pasek and Golinkoff 1996), or truth value judgment tasks (Crain and Thornton 1998) are still much more widely available and have so far provided numerous empirical data on children's language learning between birth and age five. This database will remain important and keep increasing in size: For example, the wide range of artificial language learning studies discussed by Mehler and colleagues (this volume) in the past decade was based on various behavioral experiments, and the very fact that the experimental set-up is widely available has clearly contributed to the quick growth of the literature. Furthermore, tasks like the preferential looking paradigm or the truth value judgment task, although they do not provide as precise a time course measure as ERP, are possibly more appropriate for testing grammatical constraints on sentences with multiple interpretive possibilities (Crain and Thornton 1998; Lidz, Waxman and Freedman 2003; Thornton and Wexler 1999) than ERP measures of syntactic processing, as these experiments can establish rich discourse contexts that are needed to make the relevant interpretations felicitous.

In summary, while there are obviously serious challenges in using ERP measures with children, it is very appealing that ERPs can provide much richer information about the time course and distribution of neural activities that underlie language processing, and that they allow us to compare adults and children's neu-

ral activities at various levels of linguistic processes, ranging from sounds to sentences. The addition of a new experimental technique to the field of language development is certainly a welcome one, and we believe that choosing the appropriate methodology for a given experimental hypothesis will help the field of language acquisition to further develop in an efficient fashion.

3.2   What does one stand to learn about language development from using EEG/ERP with children?

As the title of this volume suggests, our primary goal here is to improve our understanding of early *language* development by bringing together a range of behavioral and electrophysiological data. Given the cognitive science view of language as a computational system that manipulates linguistic representations, here we briefly consider whether the use of EEG/ERP with children sheds light on children's linguistic representations and their computational capacities in development.

3.2.1   *Linguistic representation*
The hallmark of linguistic research is to identify the 'parts list' at different levels of representation – that is, what are the primitives and what are their attributes? For example, phonological research suggests that the primitives are *distinctive features*, *segments*, and *syllables*. Lexical and syntactic research suggests that *morphemes* enter into linguistic computations in particular ways, and that representational elements such as *noun phrases* or *agreement markers,* and so on, constitute the representational substrates of human language. Naturally, there is vigorous debate about the nature of such representations, and it is a significant question whether research of the type presented in these chapters can speak to these critical questions about the basics of linguistic representations.

At the level of the sound structure of language, one of the basic representational units is the *phoneme* (itself made up of bundles of distinctive features), which is argued to mediate speech sounds and meaning. Although Conboy and colleagues' ERP research examines the development of *phonetic* categories that mediate speech sounds and phonological representations, it is not clear whether infants in their first two years possess *discrete phonological representations*, nor when those representations and the relevant neural mechanisms develop. From adult research (e.g. Näätänen *et al.* 1997; Phillips *et al.* 2000) we know that speakers have access to surprisingly abstract aspects of their phonological knowledge, aspects clearly not reflected in the speech signal itself. One essential representational question is whether or not learners already show evidence of such abstract categories, or whether the representations used by language learners are graded

phonetic phenomena. In the domain of sound and phonology, this is the type of question that might be engaged by innovative developmental research.

The level at which sound and meaning association is established is the lexical representation. In this domain, studies reported by Friedrich (this volume) and Sheehan and Mills (this volume) have important implications. One generalization that we can draw from their studies is that the lexical representation in the first two years of life can be noisy and require further consolidation. For example, Friedrich's finding that 12-month-olds show facilitation of phonological processing indexed by early negativity and yet lack an N400 could be interpreted to mean that the lexical representation was so noisy that the infants could not retrieve the appropriate meaning. The finding reported in Sheehan and Mills that N200–400 was elicited to known-unknown words (e.g., *bear-kobe*) as well as mispronounced known-unknown words (e.g., *gare-kobe*) in 14-month-olds but not in 20-month-olds suggests a developmental change in the 'phonetic precision' of the lexical representation between 14 and 20 months.

On the other hand, the fact that even 13 to 14-month-olds start to show N400 responses to congruous-incongruous words indicates that fairly well articulated lexical representations exist soon after the end of the first year, although we do not know how exactly the semantic memory is organized. One way to address this question is to manipulate various features of a target lexical item and test if the N400 can be modulated by how much the expected word and actual word diverge (Kutas and Federmeier 2000). If one finds that the number of features manipulated can predict differential modulation of N400 components (as in adults), this will constitute evidence for continuity in how semantic memory representations are organized from childhood to adulthood.

At the level of syntactic representation, there is a considerable amount of debate regarding whether children possess abstract structural representation early in life (see, for example, Lidz in press and Poeppel and Wexler 1993 arguing pro; Tomasello 2003 arguing con), but currently there is no contribution in this domain deriving from child ERP research. One way in which developmental ERP data could be useful to address this controversy is to find ERP evidence for syntactic priming effects (Ledoux, Traxler and Swaab 2007). If syntactic priming of, say, argument structure occurs across different verbs, this suggests that there is an abstract syntactic representation that was primed across trials. Some behavioral studies found structural priming in children (Thothathiri and Snedeker 2006) but others did not (Savage, Lieven, Theakston and Tomasello 2003). However, given that ERP is arguably more sensitive to some underlying cognitive operations than behavioral measures, there is a chance that ERP data could reveal the presence of abstract syntactic representations in the developing child grammar.

3.2.2   *Linguistic computation*

By analogy to questions regarding the elementary representational architecture of language, we would like to specify the elementary computations operating over those representations. For example, how, when and where (in the brain) are morphemes combined to yield structured representations? How are sentences built, specifically with respect to the widely attested operations such as displacement (movement), or expectancy-driven operations such as predicting the set of possible syntactic categories permissible in a given context? With regard to the first question, very little is known even in the adult literature. Although the characterizations of what might be the most basic linguistic operations must be considered one of the deepest and most pressing questions in experimental language research, we know virtually nothing about the neuronal implementation about the putative primitives of linguistic computation. On the other hand, with respect to a set of computational sub-routines that form the basis for predictive coding in language processing, both cognitive neuroscience and developmental research are making critical contributions. One nice example comes from the picture-word matching paradigms discussed by Friedrich (this volume) and Sheehan and Mills (this volume). Data deriving from studies described by these researchers are consistent with the claim that even early learners built remarkably rich representations of the on-going 'scene,' whether it is created by words, pictures or even gestures. These representations appear to make predictions about the space of possible lexical items in nuanced ways that are demonstrable using ERP responses akin to the adult N400. As such, one promising area of research at the interface between adult and child psycholinguistics might be the systematic investigation of predictive mechanisms or analysis-by-synthesis. ERP data such as ELAN and P600 (as discussed by Friederici and Oberecker, this volume) further support the hypothesis that the language processor uses subtle knowledge of language to constrain the incoming information in building meaningful representations.

3.3   Large-scale neurocognitive models – an opportunity for developmental research?

The literature on brain and language is occasionally summarized in the context of large-scale models that attempt to articulate neurocognitive frameworks for a broad range of issues in language processing. For language production, Levelt and colleagues have proposed a detailed model of single-word production (e.g., Indefrey and Levelt 2004) and Dell and colleagues have tackled sentence production (e.g., Dell, Burger and Svec 1997). With regard to comprehension, models include Hickok and Poeppel's dorsal-ventral pathway model for speech processing (e.g. Hickok and Poeppel 2007), Price's (2000) model focused on single-word process-

ing, and Friederici's (2002) model emphasizing sentence processing. Ben Shalom and Poeppel (in press) evaluate the extent to which a coherent 'meta-model' can be constructed on the basis of these differing proposals; they conclude that *storage and retrieval operations* (principally temporal lobe) versus *analytic operations* (principally parietal lobe) versus *combinatoric operations* (principally frontal lobe) provide some taxonomic structure to the issues.

While these large-scale models integrate vast amounts of data, they all share an 'adult perspective.' Almost without exception, the data considered come from studies of adult language processing – and developmental data are almost entirely absent. We view this as a major opportunity for language acquisition research with a neurobiological bent. In particular, if the large-scale heuristic models of this type characterize the adult state, it goes without saying that their developmental trajectory requires study. Friederici's (2002) model focused on sentence processing, and the developmental data presented, for example, in the Friederici and Oberecker contribution (this volume), serve to bring an explicitly developmental angle to the model. However, few of the other approaches consider language acquisition data in any depth. In our view, however, data from infant and child studies could be used productively to try to 'fractionate' the models into their constituent parts. For example, functional and functional-anatomic models arguing for a tight mapping between perception and production (e.g. Indefrey and Levelt 2004; Hickok and Poeppel 2007) should be forced to spell out how such links come to be, given typical developmental trajectories. One could imagine using behavioral and neurophysiological data to rule out (or rule in) some of the boxes that constitute the boxological models that practitioners like to entertain. The substantive addition of developmental data to such models would both test and strengthen them while dramatically increasing the models' empirical coverage.

To exemplify a further specific issue of how developmental data can shed light on the neurocognitive mechanisms underlying adults' language processing, we believe that there is another interesting opportunity to exploit in this context. There is debate in the adult electrophysiology literature on the computations reflected in the N400. The interpretations of the N400 are, by and large, rather 'high-level,' referring to processes such as semantic priming, lexical access, semantic integration into sentential contexts, and so on. But it is uncontroversial that the N400, a late and complex neurophysiological deflection, reflects various underlying neuronal computations that form the basis for these higher-order processes. This must be the case, because even a *putatively* simple cognitive operation such as lexical access is demonstrably structured and complex. Therefore it might be more fruitful to think of the (neuronally inspired) subroutines that make up processes such as lexical access or semantic priming. Given that the developmental data deal by and large with lexical processing in the absence of sentential context, these data pro-

vide a window into one of the hypothesized subroutines reflected in the N400, lexical access. As such, developmental data could be used creatively and effectively to fractionate the lexical part of the N400 into its constituent operations.

4   Recommendations from some (friendly) disciplinary neighbors

1. One missing ingredient in the field that everybody can agree on is the existence of a 'normative database' of electrophysiological responses recorded from infants and children. It goes without saying that the existence of normative criteria is all too often missing in adult studies as well. Nevertheless, the existence of a large body of research using and replicating basic response properties allows researchers to compare across studies, not just with respect to the issues under investigation, but specifically with respect to the response profile obtained across studies. This may include the latency and amplitude and spatial distribution of responses, the effect size of responses in a given study, and so on.

As Table 1 illustrates, even the relatively small selection of papers illustrated in the present volume shows the heterogeneity of responses associated with linguistic stimulation. But would a more extensive database be more than a list of vaguely associated phenomena? Why might such a database even be illuminating? Consider from the present volume three papers that use the N400 response to test aspects of lexical semantic processing. Conboy and colleagues (this volume) describe experimental research in which the N400 is elicited in sentential contexts. This is akin to many standard designs in adult psycholinguistic research. In contrast, Friedrich (this volume), as well as Sheehan and Mills (this volume), uses a picture-word matching paradigm and demonstrates N400 modulation. Now, the presupposition underlying this work must be that the cognitive operation or set of operations underlying the generation of the N400 must be to some extent shared across these studies. In other words, although the 'lead-in' processes are quite different in a sentence-processing versus a picture matching paradigm, it stands to reason that in both cases, a specific expectation is built up for a lexical semantic representation, and perhaps even a specific lexical entry. If the N400 reflects aspects of retrieving lexical representations, then the N400 pattern—at least if we look at it within an age cohort—should look rather comparable. Obviously this can only be verified with respect to some larger database in which the many conditions that elicit the N400 are tabulated and compared in a quantitative manner. In order to build a theoretically grounded and computationally satisfying account of what is reflected in N400 responses, it seems to us that such a database would be of immense value. The necessity of the normative database is very clear and urgent,

given that different researchers are already finding these diverse components and assigning variable interpretations to them.

A further feature of a normative database that would be extremely valuable, although not tremendously amusing to establish, would be to have extensive data for given well-known responses across development. For example, the ontogenetically conditioned modification of the N1 is known, as is the fact that the MMN response can be elicited throughout development. On the other hand, the timing, amplitude, and spatial profile of, say, the N2, or N400, or P600 is not at all understood. Some laboratory will earn lots of praise for collecting, for example, N400 and P600 responses from the cradle to the grave.

2. Because of the extraordinary variability of electrophysiological responses as a function of age, the interpretation of these components vis-à-vis cognitive models is even more challenging. One small but potentially useful modification might be the following: If, in the context of practically all experiments, a 'functional landmark' was established, then the data could be interpreted with respect to the robust landmark that in turn should be more easily comparable across studies. What we have in mind is roughly the following: Suppose that at the beginning (or middle, or end) of a psycholinguistic study a simple evoked response to a supra-threshold tone was recorded (for example, N1), then it would be possible to interpret the psycholinguistically elicited ERP responses with respect to that more well understood functional landmark. It is well known that the N1 itself undergoes a rather dramatic change from infancy to late puberty (Pang and Taylor 2000; Wunderlich and Cone-Wesson 2006), but within age groups, it would be possible to compare the responses, which would now be scaled to something rather more straightforwardly interpretable. One reason we raise this possibility is the extreme variability in timing and amplitude of responses. If it were possible to interpret, say, an N400 pattern as a relative difference to a simple tone instead of an absolute time, this opportunity would constrain the interpretation one would give language-related ERPs.

3. As is true for adult studies, one feature of such research that would enrich our understanding, both with respect to linguistic and with respect to neurobiological interpretation, is the 'radical' decomposition of the tasks used with infants and children into their constituent psycholinguistic and cognitive subroutines. It is entirely uncontroversial that an experimental task such as lexical decision or picture-word matching encompasses numerous, more elementary computational subroutines. While we tend to interpret evoked responses as somewhat monolithic (the N400 reflects "semantic integration"; the P600 reflects "repair and reanalysis"), researchers agree that these are not the elementary operations executed by neuronal circuitry. Rather, many different operations go into the composition of these complicated brain responses. A richer, more satisfying account of develop-

ment will presumably give an account of the individual operations that make up, say, lexical access or semantic composition.

5   Conclusion

Even brief historical reflection should convince the reader that it is remarkable that we are now able to record extra-cranial signals from young brains with milli-second resolution, and that these signals *can* receive a coherent interpretation in the context of theories about language development and language processing. Having solved some of the major technical obstacles, the question is now whether harnessing these new techniques and optimizing them to investigate child language will yield knowledge that goes substantially beyond what we can learn from sophisticated behavioral techniques. In our view, it is terrific news when the behavioral data and the ERP data match and correlate – in some sense that can be viewed as replication with different methods. However, while it is nice when behavior and physiology match, it is also boring. When there exists substantial parallelism between ERP data and behavioral data, what is added? The experimental situation is at its most interesting when the two types of data go against each other or complement each other. It is precisely in those cases that one gets the sense of new facets of data being available to answer questions about development at a fine-grained neurobiological and cognitive level.

Finally, there is one aspect of language acquisition research in which the developmental literature is empirically more broadly informed than most of the adult literature, and that concerns the incorporation of a cross-linguistic perspective. An explicitly multilingual angle is dictated by the need to account for cross-linguistic facts about language development. Adult cognitive neuroscience of language research could benefit from this linguistic pluralism.

Acknowledgments

References

Ben Shalom, D. and Poeppel, D. In press. "Functional anatomic models of language: Assembling the pieces." *The Neuroscientist*.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

Crain, S. and Thornton, R. 1998. *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: The MIT Press.

Dell, G.S., Burger, L.K. and Svec, W.R. 1997. "Language production and serial order: A functional analysis and a model." *Psychological Review* 104: 123–47.

Elman, J.L. 1993. "Learning and development in neural networks: The importance of starting small." *Cognition* 48: 71–99.

Friederici, A.D. 2002. "Towards a neural basis of auditory sentence processing." *Trends in Cognitive Science* 6: 78–84.

Friederici, A.D. 2005. "Neurophysiological markers of early language acquisition: from syllables to sentences." *Trends in Cognitive Science* 9 (10): 481–8.

Friederici, A.D. and Weissenborn, J. 2007. "Mapping sentence form onto meaning: The syntax-semantic interface." *Brain Research* 1146: 50–8.

Hickok, G. and Poeppel, D. 2007. "The cortical organization of speech perception." *Nature Neuroscience Reviews* 8: 393–402.

Hirsh-Pasek, K. and Golinkoff, R.M. 1996. *The origins of grammar*. Cambridge, MA: The MIT Press.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R. and Cutler, A. 2000. "Crosslanguage word segmentation by nine-month-olds." *Psychonomic Bulletin and Review* 7: 504–9.

Indefrey, P. and Levelt, W.J. 2004. "The spatial and temporal signatures of word production components." *Cognition* 92: 101–44.

Johnson, M.H. 2005. *Developmental cognitive neuroscience,* 2nd edn. Malden, MA: Blackwell.

Jusczyk, P.W. 1997. *The discovery of spoken language*. Cambridge, MA: The MIT Press.

Jusczyk, P.W. and Aslin, R.N. 1995. "Infants' detection of the sound patterns of words in fluent speech." *Cognitive Psychology* 29: 1–23.

Kemler-Nelson, D. G., Jusczyk, P.W., Mandel, D.R., Myers, J. and Turk, A.E. 1995. "The Head-turn Preference Procedure for testing auditory perception." *Infant Behavior and Development* 18: 111–16.

Kuhl, P. 2004. "Early language acquisition: Cracking the speech code." *Nature Reviews Neuroscience* 5: 831–43.

Kutas, M. and Federmeier, K.D. 2000. "Electrophysiology reveals semantic memory use in language comprehension." *Trends in Cognitive Science* 4: 463–70.

Ledoux, K., Traxler, M.J. and Swaab, T.Y. 2007. "Syntactic priming in comprehension: Evidence from event-related potentials." *Psychological Science* 18 (2): 135–43.

Lidz, J. In press. "The abstract nature of syntactic representations: Consequences for a theory of learning." In *Handbook of language development*, E. Hoff and M. Shatz (eds.), Cambridge: Blackwell.

Lidz, J., Waxman, S. and Freedman, J. 2003. "What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18-months." *Cognition* 89: B65-B73.

Marcus, G. F., Vijayan, S., Rao, S.B. and Vishton, P. 1999. "Rule learning by seven-month-old infants." *Science* 283 (5398): 77–80.

Näätänen, R., Lehtoskoskl, A., Lennes, M., Cheour, M., Huotilainen, M., Ilvonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J. and Alho, K. 1997. "Language-specific phoneme representations revealed by electric and magnetic brain responses." *Nature* 385: 432–34.

Nelson, C.A. and Luciana, M. (eds.). 2001. *Handbook of developmental cognitive neuroscience.* Cambridge, MA: The MIT Press.

Newport, E. and Aslin, R. 2000. "Innately constrained learning: Blending old and new approaches to language acquisition." In *Proceedings of the 24th Annual Boston University Conference on Language Development,* S.C. Howell, S.A. Fish and T. Keith-Lucas (eds.), Somerville, MA: Cascadilla.

Pang, E.W. and Taylor M.J. 2000. "Tracking the development of the N1 from age 3 to adulthood: An examination of speech and non-speech stimuli." *Clinical Neurophysiology* 111 (3): 388–97.

Pasman, J.W., Rotteveel, J.J., Maassen, B. and Visco, Y.M. 1999. "The maturation of auditory cortical evoked responses between (preterm) birth and 14 years of age." *European Journal of Paediatric Neurology* 3: 79–82.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M. and Roberts, T.P. 2000. "Auditory cortex accesses phonological categories: An MEG mismatch study." *Journal of Cognitive Neuroscience* 12 (6): 1038–55.

Phillips, C. and Wagers, M. In press. "Relating structure and time in linguistics and psycholinguistics." In *Oxford handbook of psycholinguistics*, G. Gaskell (ed.). Oxford: OUP.

Pinker, S. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Poeppel, D. and Wexler, K. 1993. "The full competence hypothesis of clause structure in early German." *Language* 69: 1–33.

Poeppel, D. and Embick, D. 2005. "Defining the relation between linguistics and neuroscience." In *Twenty-first century psycholinguistics: Four cornerstones*, A. Cutler (ed.), 103–18. Mahwah, NJ: Lawrence Erlbaum Associates.

Poeppel, D., Idsardi, W.J. and van Wassenhove, V. In press. "Speech perception at the interface of neurobiology and linguistics." *Philosophical Transactions of the Royal Society*.

Price, C.J. 2000. "The anatomy of language: contributions from functional neuroimaging." *Journal of Anatomy* 197 (Pt 3): 335–59.

Saffran, J., Aslin, R. and Newport, E. 1996. "Statistical learning by 8-month- old infants." *Science* 274 (5294): 1926–8.

Savage, C., Lieven, E., Theakston, A. and Tomasello, M. 2003. "Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children." *Developmental Science* 6: 557–67.

Thornton, R. and Wexler, K. 1999. *Principle B, VP-ellipsis, and interpretation in child grammar*. Cambridge, MA: The MIT Press.

Thothathiri, M. and Snedeker, J. 2006. "Children's representation of verbs: Evidence from priming during online sentence comprehension." In *Proceedings of the 30th Annual Boston University Conference on Language Development*, D. Bamman, M. Tatiana and C. Zaller (eds.), 643–652. Somerville, MA: Cascadilla.

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

Werker, J.F. and Tees, R. C. 1999. "Influences on infant speech processing: Toward a new synthesis." *Annual Review of Psychology* 50: 509–35.

Woodward, J.Z. and Aslin, R.N. 1990. "Segmentation cues in maternal speech to infants." Paper presented at 7th biennial meeting of the International Conference on Infant Studies, Montreal, Quebec, Canada.

Wunderlich. J.L. and Cone-Wesson B.K. 2006. "Maturation of CAEP in infants and children: a review." *Hearing Research* 212 (1–2): 212–23.

Yuille, A. and Kersten, D. 2006. "Vision as Bayesian inference?: analysis by synthesis?" *Trends in Cognitive Science* 10 (7): 301–8.