Brain Research

MAY 7, 2008 | VOLUME 1208

ISSN: 0006-8993

Research Report

A developmental ERP study of verbal and non-verbal semantic processing

Alycia Cummings[a,b,c,*], Rita Čeponienė[a], Frederic Dick[c,e],
Ayse Pinar Saygin[c,f], Jeanne Townsend[a,d]

[a]Project in Cognitive and Neural Development, University of California, San Diego, USA
[b]San Diego State University/University of California, San Diego Joint Doctoral Program in Language and Communicative Disorders, USA
[c]Center for Research and Language, 9500 Gilman Drive, University of California, San Diego, UCSD Mail Code 0526,
La Jolla, CA 92093-0526, USA
[d]Department of Neurosciences, University of California, San Diego, USA
[e]Birkbeck College, University of London, UK
[f]Institute of Cognitive Neuroscience and Functional Imaging Laboratory, University College London, UK

ARTICLE INFO

ABSTRACT

To clarify how different the processing of verbal information is from the processing of meaningful non-verbal information, the present study characterized the developmental changes in neural responses to words and environmental sounds from pre-adolescence (7–9 years) through adolescence (12–14 years) to adulthood (18–25 years). Children and adults' behavioral and electrophysiological responses (the N400 effect of event-related potentials) were compared during the processing of words and environmental sounds presented in semantically matching and mismatching picture contexts. Behavioral accuracy of picture–sound matching improved until adulthood, while reaction time measures leveled out by age 12. No major electrophysiological changes in the N400 effect were observed between pre-adolescence and adolescence. When compared to adults, children demonstrated significant maturational changes including longer latencies and larger amplitudes of the N400 effect. Interestingly, these developmental differences were driven by stimulus type: the Environmental Sound N400 effect decreased in latency from adolescence to adulthood, while no age effects were observed in response to Words. Thus, while the semantic processing of single words is well established by 7 years of age, the processing of environmental sounds continues to improve throughout development.

1. Introduction

Mechanisms underlying the development of language comprehension are still not fully understood. One account postulates that language skills develop on the basis of innate, 'special' processing resources that are devoted to linguistic, symbolic, and grammatical processes (e.g. Pinker, 1994). Alternatively, language may emerge from, and along with, non-linguistic skills that provide component "building blocks" for language abilities, such as sensory, motor, joint attention, and gesturing abilities

(Thelen and Smith, 1994; Dick et al., 2004; MacWhinney, 1999). In the latter case, the processing of meaningful information could originate in a common, domain-general system and through development could be refined into separate categories of information. One way to test whether or not the processing of linguistic and non-linguistic information follows the same path during development is to compare behavioral and neural responses to spoken words with those to environmental sounds.

1.1. Processing of words vs. environmental sounds

Speech and environmental sounds are two different types of auditory information that can serve the same purpose: they convey meaningful information involving people and environmental events. Thus, our perception and cognition of both types of auditory information may proceed in a similar fashion. Both speech and environmental sounds are acoustically rich and complex, and the processing of both breaks down with acoustic degradation (Gygi, 2001; Shafiro and Gygi, 2004). The processing of environmental sounds is enhanced by context and item familiarity (Ballas and Howard, 1987; Ballas, 1993; Cycowicz and Friedman, 1998). However, environmental sounds also differ from speech along some important dimensions. There is wide individual variation in exposure to different environmental sounds (Gygi, 2001). Thus, healthy adults show much variability in their ability to recognize and identify these sounds (Saygin et al., 2005).

Recent behavioral evidence suggests that word and sound processing change similarly during infancy (Cummings et al., submitted), typical development (Borovsky et al., submitted), aging (Saygin et al., 2005), and in brain lesion patients with aphasia (Saygin et al., 2003). Although word and meaningful non-linguistic sound processing yields similar behavioral indices, the neural processing routes involved in verbal vs. meaningful non-nonverbal processing might be different. One measure that allows identification and assessment of distinct, well-defined stages of meaningful auditory processing is time-sensitive event-related potentials (ERPs).

1.2. The N400 effect

Event-related potentials (ERPs) can identify the rapid succession of neural processing stages and coarse configurations of the neural networks associated with different information processing types. One of the ERP components closely tied to semantic processing is a negative wave peaking at approximately 400 ms post-stimulus onset, the N400 (Kutas and Hillyard, 1980a,b; Kutas and Hillyard, 1983). All semantic stimuli (auditory or visual, orthographic or pictorial) elicit an N400, but its amplitude and latency vary depending on the preceding context: the N400 is larger in amplitude when the stimulus does not match an expectancy set by the fore-going message. Importantly, when put into a similar context-dependent situation, environmental sounds elicit N400 peaks that are similar to those elicited by auditory or visual words (Van Petten and Rheinfelder, 1995; Plante et al., 2000).

Using a cross-modal sound-picture match/mismatch paradigm, we have compared brain responses to words and environmental sounds in a group of young adult subjects (Cummings et al., 2006). In this paradigm, environmental sound and word trials were made maximally comparable by equating for dura-

tion, amplitude, and norming-verified semantic content. No N400 effect amplitude differences between the two types of stimuli were found. Interestingly, the word-elicited N400 effect peaked later ($M = 401$ ms) than the environmental sound-elicited N400 effect ($M = 331$), even though there were no behavioral reaction time differences. These results indicated that initially, environmental sound processing might proceed at a faster rate than word processing because the former do not need to undergo lexical encoding. Instead, they can be directly mapped onto their semantic representations. However, transforming semantic identification into a behavioral response may take a relatively longer period of time for the environmental sounds than words. To a point, this may be an experiential effect, given that an average adult has not only more exposure to verbal information but also generates speech for communication.

1.3. The development of the N400

One of the first large-scale ($n = 130$) developmental N400 ERP studies looked into semantic processing during auditory and visual sentence comprehension (Holcomb et al., 1992) in children and young adults between the ages of 5 and 26. Subjects were presented with visual and auditory sentences that ended with either highly expected or semantically inappropriate words. Marked differences were revealed in the N400 effect, with longer latency and increased amplitude from 5 to 16 years of age, after which they stabilized at adult-like levels. Additionally, unlike adults who typically display a Right>Left asymmetry when processing anomalous words (e.g. Neville et al., 1986), robust hemispheric asymmetries were not observed in children until 13 years of age. Holcomb et al. (1992) suggested that considerable cerebral reorganization throughout development led to the increasing functional specificity and decreased reaction times observed in their subjects.

The results of Holcomb et al. (1992) were essentially replicated in a Finnish study by Juottonen et al. (1996). In that study, children (aged 5–11 years) and adults heard semantically congruous and incongruous sentences. The N400 effect was observed to be much larger in amplitude and longer in latency in children than adults. Additionally, the N400 was maximal over parietal sites in children, while the adult N400 effect was centered slightly anterior to that of the children. Thus, Juttonen et al. (1996) also concluded that the N400 semantic context effect declined with age.

Another ERP study examined children's (aged 6–13 years) ability to detect semantic violations using German sentences (Hahne et al., 2004). The latency of the N400 was longest in 6- to 8-year-olds, and became similar to that of adults by age 10. However, unlike in the Holcomb et al. (1992) study, a decrease in N400 amplitude with age was not found. Based on their results, Hahne et al. (2004) argued that the neurophysiological basis for sentence-level semantic processing does not drastically change between early and late childhood. Alternatively, this result may have been due to the fact that all the children above 7 years of age had similar error rates (only the 6-year-olds had significantly higher error rates). In other words, the stimuli may not have been difficult enough to tease apart developmental effects.

To our knowledge, no study has examined children's processing of meaningful non-verbal (environmental) sounds alone or in comparison with corresponding verbal items.

Table 1 – Accuracy and Median Reaction Time (SD) measures for Words and Environmental Sounds for Pre-Adolescents, Adolescents, and Adults

Sound type	All groups		Pre-adolescents		Adolescents		Adults	
	Accuracy % correct	RT in ms	Accuracy % correct	RT in ms	Accuracy % correct	RT in ms	Accuracy % correct	RT in ms
Words	91.63 (7.73)	870 (213)	85.90 (8.59)	1079 (210)	92.81 (7.01)	777 (146)	96.42 (3.04)	782 (134)
Environmental sounds	93.71 (6.26)	886 (207)	88.03 (6.51)	1078 (176)	93.70 (4.78)	808 (153)	98.64 (1.68)	798 (167)
Both sound types			86.97 (7.62)	1079 (192)	92.76 (6.03)	792 (149)	97.53 (2.68)	790 (150)

Measures are pooled across Noun and Verb experiments.

Understanding how the brain processes verbal information in comparison to other types of meaningful information is important for understanding the neural basis of language. More specifically, developmental studies using such a design could elucidate the nature and the timeline of lexical specialization. Such a comparison done in children and adults could demonstrate whether or not processing of linguistic information undergoes specialization during development — that is, whether or not it diverges from processing of meaningful non-verbal information.

Instead of directly contrasting the domain specific and domain-general theories of semantic processing, the current study focused on examining the differences that may or may not be present during the processing of words as compared to environmental sounds in children and adults. It is possible that at the very early stages of sensory processing, all acoustic information is processed by similar mechanisms, while at a later stage of processing, the verbal and environmental sound information may converge onto the same high-level semantic representation. However, there may be intermediate stages during which processing of environmental sounds and words diverge. For example, one type of auditory input may lend itself to more automatic processing (i.e., faster encoding) than the other. It was the aim of the present study to identify and characterize those differences. We hypothesize that both similarities and differences exist during the processing of environmental sounds and words and the goal of the current study was to clarify how those differences and similarities are represented in the electrophysiological brain activity of children and adults.

While the task used in the present study has been validated in a group of healthy adults, as described in the above-mentioned study by Cummings et al. (2006), the previous adult study and the present developmental study had different goals. The goal of the adult study (Cummings et al., 2006) was

to compare the verbal and non-verbal processing in the mature brain, as well as to further explore the nature of the properties of the N400 effect and related brain components. The aim of the present study was to examine developmental changes (if any) in the behavioral and ERP indices established in the adult study.

In summary, the goals of the current study were as follows: 1) To compare behavioral and electrophysiological measures of semantic integration of words and environmental sounds in children; and 2) To examine developmental changes in the respective neural responses, the word and environmental sound N400 effects, across groups of Pre-Adolescent children (7–9 years of age), Adolescent children (12–14 years of age), and Adults (18–25 years of age).

2. Results

2.1. Accuracy

Overall, participants responded more accurately to Environmental Sounds than to Words ($F(1,40) = 31.19$, $p < .0001$; Table 1). A main effect of Age Group was observed ($F(2,40) = 19.05$, $p < .0001$). Pre-planned contrasts revealed that accuracy improved for each age group: Adolescents were more accurate than the Pre-adolescents ($F(1,40) = 11.43$, $p < .002$) and Adults were more accurate than Adolescents ($F(1,40) = 8.39$, $p < .007$). No Sound Type × Age Group interaction was observed ($F(2,40) = .07$, $p > .92$).

Mismatching trials also elicited greater accuracy ($M = 94.04\%$) than Matching trials ($M = 91.3\%$; $F(1,40) = 17.63$, $p < .0001$). Further, a strong Sound Type (Word/Environmental Sound) × Trial Type (Match/Mismatch) interaction was found ($F(1,40) = 31.11$, $p < .0001$). Post-hoc ANOVAs revealed that while responses to the Environmental Sounds were not affected by Trial Type ($M =$

Table 2 – N400 Latency (SD) in milliseconds for Pre-Adolescents, Adolescents, and Adults recorded at the midline electrodes

Sound type	All groups	Pre-adolescent				Adolescent				Adult			
		OV	Fz	Cz	Pz	OV	Fz	Cz	Pz	OV	Fz	Cz	Pz
Words	394 (31)	392 (28)	394 (33)	384 (23)	379 (23)	389 (23)	390 (22)	389 (27)	387 (27)	399 (39)	414 (41)	397 (32)	387 (36)
Environmental sounds	360 (41)	386 (25)	388 (28)	383 (25)	380 (21)	374 (24)	374 (26)	369 (22)	368 (22)	325 (41)	317 (44)	325 (43)	326 (43)
Both sound types		389 (27)				381 (25)				362 (55)			

Measures are pooled across Noun and Verb experiments.
*OV = overall.

93.61% and 93.82% for Match and Mismatch trials, respectively), subjects' responses to the Words were less accurate in Matching than Mismatching trials (M=89% and 94.26%, respectively; $F(1,40)=36.23$, $p<.0001$). In other words, subjects had more difficulty determining whether or not a word matched (described) an object than determining that it did not match. However, no such dependency was found for the Environmental Sounds. Trial Type did not interact with the Age Group ($F(2,40)=1.18$, $p>.31$).

2.2. Reaction Time

Sound Type (Word/Environmental Sound) did not affect reaction times ($F(2,40)=1.71$, $p>.19$; Table 1). A main effect of Age Group was observed ($F(2,40)=14.87$, $p<.0001$). However, this effect was driven by the slower reaction times of the Pre-adolescent children, which were significantly slower than both the Adolescents ($F(1,40)=22.63$, $p<.0001$) and Adults ($F(1,40)=23.03$, $p<.0001$). The two older groups did not significantly differ from each other

WORDS

ENVIRONMENTAL SOUNDS

— Pre-adolescents (7-9 years)
— Adolescents (12-14 years)
···· Adults (18-25 years)
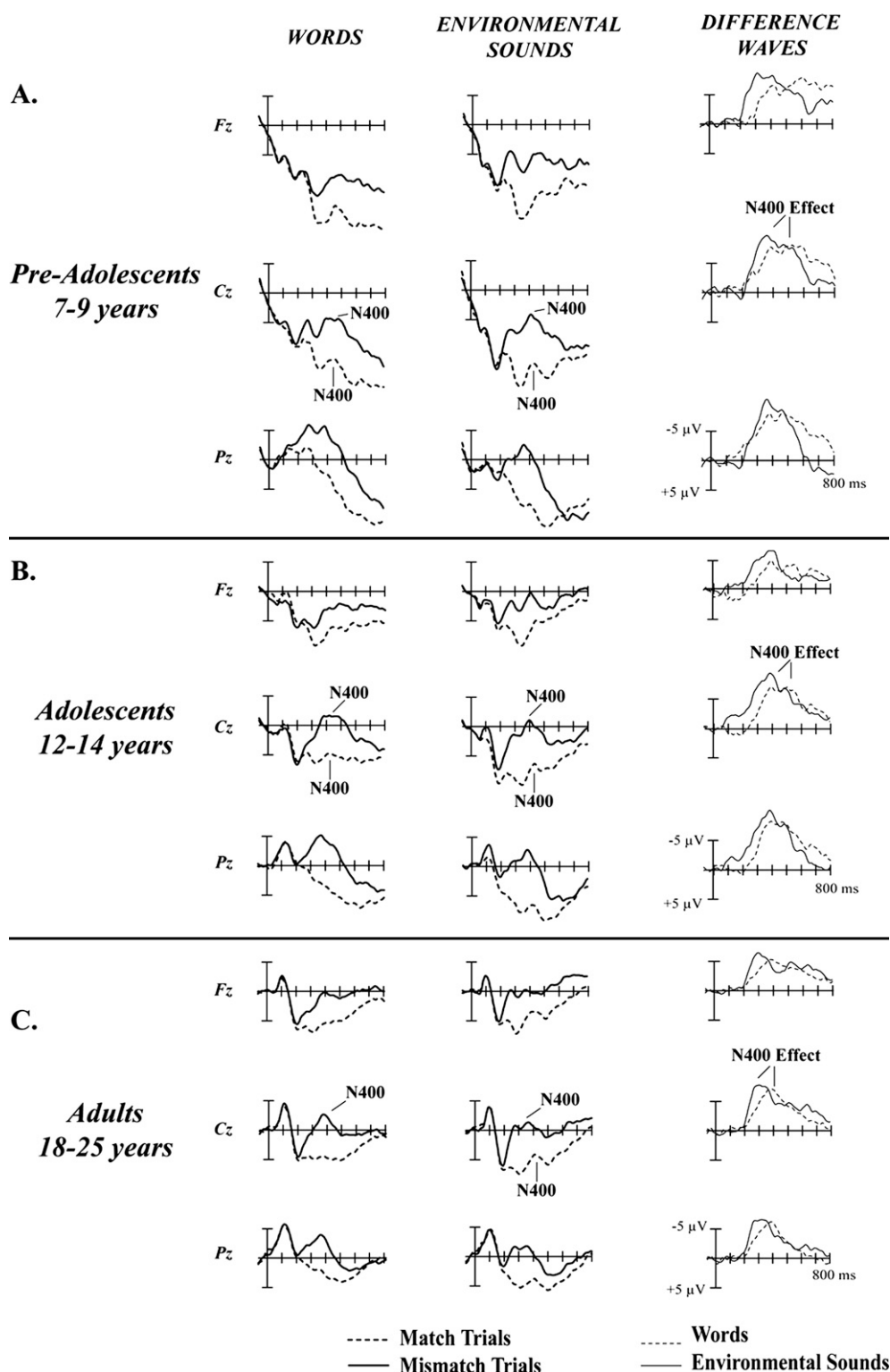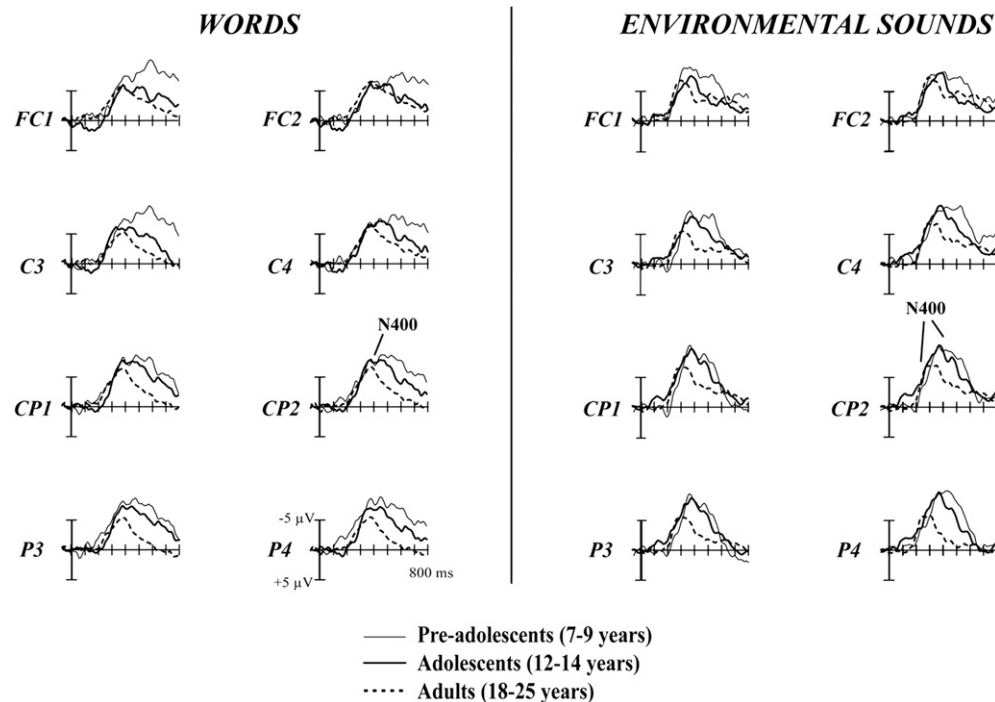
Fig. 2 – Mismatch-ERP minus Match-ERP difference waves to Words and Environmental Sounds for the Pre-adolescents, Adolescents, and Adults. All three age groups elicited similar N400 amplitudes and latencies in response to the Word stimuli. In response to the Environmental Sound stimuli, the Adults elicited much earlier N400 effect responses than did the two groups of children. The Adult N400 effect was also significantly smaller than that of the Pre-Adolescent children.

($F(1,40)=.002$, $p>.96$). No Sound Type×Age Group interaction was found ($F(2,40)=.64$, $p>.53$). In contrast to accuracy data, reaction times were faster in Matching than Mismatching trials ($M=867$ ms vs. 889 ms; $F(1,40)=6.89$, $p<.02$). Again, no interactions involving Age Group were found ($F(2,40)=.84$, $p>.43$).

2.3. N400 effect latency

Environmental Sound stimuli elicited an earlier N400 effect than Word stimuli ($F(1,40)=34.07$, $p<.0001$; Table 2, Fig. 1). Age Group was also a significant effect ($F(2,40)=11.11$, $p<.0001$). Pre-planned contrasts revealed that while the Pre-adolescents' and Adolescents' N400 effect peak latencies did not differ ($F(1,40)=1.60$, $p>.21$), both groups demonstrated significantly later N400 effects than the Adults (Pre-adolescents: $F(1,40)=20.36$, $p<.0001$; Adolescents ($F(1,40)=11.35$, $p<.002$; Fig. 2).

Interestingly, an Age Group×Sound Type (Word/Environmental Sound) interaction was also found for the N400 effect latency data ($F(2,40)=15.68$, $p<.0001$; Fig. 2). Post-hoc ANOVAs revealed that the Word N400 effect latency did not differ across Age Groups ($F(2,40)=.79$, $p>.46$) while the Environmental Sound N400 effect latency decreased from childhood to adulthood ($F(2,40)=25.84$, $p<.0001$). Specifically, while the Pre-adolescents and Adolescents did not differ from one another ($F(1,40)=1.76$, $p>.19$), the Adult Environmental Sound N400 effect was significantly earlier than that of the Pre-Adolescent ($F(1,40)=44.59$, $p<.0001$) and the Adolescent ($F(1,40)=30.82$, $p<.0001$) N400 effects.

2.4. N400 effect amplitude

Environmental Sound stimuli elicited a significantly larger N400 effect than Word stimuli ($F(1,40)=5.19$, $p<.03$; Table 3, Figs. 1 and 2). A strong trend for the main Age Group effect was also observed ($F(2,40)=3.05$, $p<.06$). Pre-planned contrasts revealed no difference between the Pre-adolescents' and Adolescents' N400 effects ($F(1,40)=.34$, $p>.56$). However, the Adult N400 effect was significantly smaller in amplitude than the Pre-

Fig. 1 – Matching and Mismatching Word (left column) and Environmental Sound (middle column) ERPs and Difference Waves (right column). A. Pre-Adolescents. The N400 component is clearly visible in the matching and mismatching ERPs in response to words and environmental sounds in the 300–500 ms range. The N400 effect is a difference between the matching and mismatching ERPs in response to words and environmental sounds in the 300–500 ms range. The N400 effect appears as a negative peak in the mismatch-ERP minus match-ERP difference waves. B. Adolescents. The N400 component is clearly visible in the matching and mismatching ERPs in response to environmental sounds and mismatching words, though diminished in response to matching word trials. The N400 effect is clearly visible as a negative peak in the difference waves. C. Adults. The N400 component is clearly visible in the matching and mismatching ERPs in response to environmental sounds and mismatching ERPs to words, but barely visible in response to matching word trials. The N400 effect is clearly visible as a negative peak in the difference waves.

Table 3 – N400 Amplitude (SEM) negative microvolts for Pre-Adolescents, Adolescents, and Adults recorded at three midline electrodes

Sound type	All Groups	Pre-adolescent				Adolescent				Adult			
		OV	Fz	Cz	Pz	OV	Fz	Cz	Pz	OV	Fz	Cz	Pz
Words	6.13	6.88	6.91	7.94	8.70	6.31	4.99	7.68	9.06	5.30	6.02	7.98	7.01
	(.29)	(.29)	(1.56)	(1.61)	(1.41)	(.20)	(.75)	(1.32)	(1.41)	(.16)	(.71)	(.94)	(.84)
Environmental sounds	7.26	8.30	9.00	9.99	10.31	7.85	7.42	9.86	10.26	5.77	8.67	8.62	7.37
	(.13)	(.26)	(1.30)	(1.04)	(1.40)	(.23)	(1.09)	(.97)	(1.02)	(.18)	(.94)	(.97)	(.75)
Both sound types		7.59 (.20)				7.08 (.15)				5.53 (.12)			

Measures pooled across Noun and Verb experiments. All mean amplitudes were significant at $p < .0001$ as compared to pre-stimulus baseline.
*OV = overall.

Adolescent N400 effect ($F(1,40) = 5.46$, $p < .03$) and was nearly significantly smaller than the Adolescent N400 effect ($F(1,40) = 3.33$, $p < .08$). No Sound Type × Age Group interaction was observed ($F(2,40) = .47$, $p > .62$).

2.5.   N400 *effect scalp distribution*

A main effect of Anterior-Posterior distribution was observed ($F(6,240) = 15.16$, $p < .0001$). Electrodes at the most anterior (F3/F4) and posterior (O1/O2) scalp sites elicited the smallest activation difference ($M = -5.91$ μV and -5.26 μV, respectively), while the greatest activation difference was elicited over the centro-parietal electrode sites (CP1/CP2; $M = -8.65$ μV; Fig. 2). However, no Sound Type × Anterior–Posterior ($F(6,240) = 1.79$, $p > .10$) or Age Group × Anterior-Posterior ($F(12,240) = 1.15$, $p > .31$) interaction was found using normalized N400 effect amplitudes.

A main Laterality effect was also observed ($F(1,40) = 22.42$, $p < .0001$). Both Words and Environmental Sounds elicited greater activation over the Right ($M = -5.23$ μV and -6.1 μV, respectively) than the Left hemisphere ($M = -3.99$ μV and -4.59 μV, respectively). No Sound Type × Laterality ($F(1,40) = .02$, $p > .88$) or Age Group × Laterality ($F(2,40) = .04$, $p > .96$) interaction was found. Thus, similar scalp distributions of Words and Environmental Sounds were observed across all three Age Groups.

3.   Discussion

This study compared behavioral and electrophysiological indices associated with semantic processing of Words and Environmental Sounds in Pre-adolescent children (7–9 years), Adolescent children (12–14 years), and Adults (18–25 years). Behavioral response accuracy improved with each older age group, while reaction times reached adult-like values by 12 years of age. Few differences were observed in how meaningful verbal and non-verbal information was processed in children. While previous ERP studies (e.g., Holcomb et al., 1992) reported developmental differences through the age of 16, the present study found that the age groups demonstrating significant developmental effects in brain electrophysiology were Adolescents versus Adults, and not the two younger groups. Moreover, developmental changes were not consistent across the verbal and non-verbal meaningful stimuli. Namely, the latency of the Environmental Sound N400 effect shortened between Adolescence (12–14 years) and Adulthood. No age-related changes in the Word N400 effect were found.

3.1.   Sound type (Word/Environmental sound) effects

Subtle though consistent differences between Word and Environmental Sound processing were observed in the present study. Behavioral accuracy was higher with the Environmental Sounds than Words, while no behavioral reaction time differences between Words and Environmental Sounds were observed. The higher behavioral accuracy to Environmental Sounds over Words paralleled the larger N400 effect to Environmental Sounds than Words. However, the reaction time data was discordant with the latency of the N400 effect, which was shorter for the Environmental Sounds than Words.

One explanation for the finding of the larger Environmental Sound N400 effect could be a relative salience of sound–object compared to word–object association. More specifically, fewer semantic entities of the environmental sounds, as compared to words, can be mapped onto any given object or action.[1] In contrast, many different words (nouns and verbs) can be associated with the same object. Moreover, words may be the more expected or familiar label for simple pictures depicting objects and actions. Such an unequal association could have made the match/mismatch task more complex specifically in the Word trials. If so, any picture would be associated with, and pre-activate, a greater number of "average" Word representations than Environmental Sound representations. This may have diminished the N400 effect and prolonged its latency in the Word as compared with the Environmental Sound trials.

A complimentary interpretation takes into account the much higher frequency of the co-occurrence of objects/actions and their associated environmental sounds, as compared to word labels. For example, when a helicopter is seen flying overhead, invariably the sound is heard. It does not take many exposures to this object–sound pairing before just the sight of the object will elicit expectations of its associated sound. On the other hand, someone labeling and object/action does not occur with such systematic regularity (e.g., "Look, a helicopter is whirling overhead"). As a result, the expectations of hearing an associated environmental sound when seeing an object/action would result in a larger N400 effect, reflecting stronger

—————————

[1] Semantically, not acoustically.
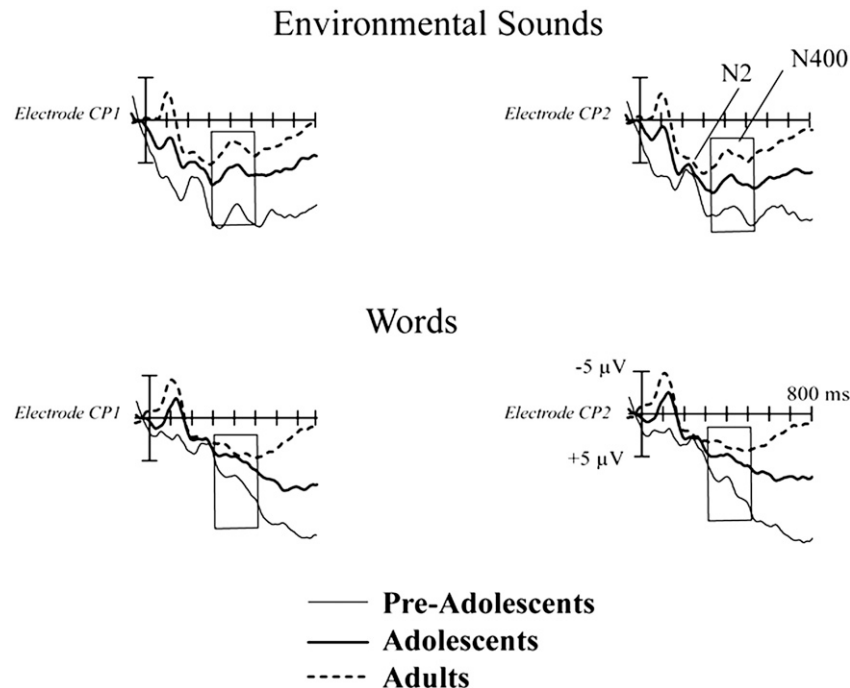
Fig. 3 – Matching ERP responses to Word and Environmental Sound trials in Pre-Adolescents, Adolescents, and Adults. The matching-item N400 was elicited, and was of a strikingly similar amplitude and appearance, in all three age groups in the Environmental Sound trials (top panel). N400 was much less pronounced in the Word trials (bottom panel). The three main developmental effects were: (1) the smaller N1, (2) the larger N2 (marked), and (3) the larger slow underlying positive shift activity in the youngest groups. These N1 and N2 effects were outside the N400 analysis window.

semantic expectations that are violated. Additionally, the greater amount of expectancy of an associated environmental sound would also lead to shorter N400 effect latencies since the expectancies would be more systematically present upon seeing an object/action.

To examine this possibility further, we overlaid the three subject groups' matching trial responses to the words and environmental sounds (Fig. 3). The matching-item N400 elicited by the environmental sounds was strikingly similar in amplitude and appearance across all three age groups. On the other hand, the N400 elicited by the words was poorly expressed in all age groups. Thus, these results could be interpreted as a result of the relative salience or higher frequency of co-occurrence of objects/actions and their associated environmental sounds as compared to words.

Alternatively, it is possible that the verbal nature of words requires that they go through an additional (lexical) stage of processing before their semantic nature can be recognized, while environmental sounds can be semantically interpreted based directly on their acoustic patterns. This would afford a temporal advantage to Environmental Sounds and lead to the earlier N400 effect peak latency.

On the surface, the fact that subjects' reaction times (via button press) were similar to Words and Environmental sounds appears discordant to the N400 effect latency differences. As argued in Cummings et al. (2006), this discrepancy might originate in the "output" stage of processing. While initially, an environmental sound may be more quickly identified, the present data suggests that turning that identification into a behavioral response takes a longer period of time for Environmental Sounds than for Words. It is possible that word stimuli have more efficient associations with various response mechanisms than do environmental sounds due to the daily production of words and an ongoing adjustment of behavior in response to verbal input. Thus, translating the identification of a meaningful, though non-verbal, sound into a behavioral response may take longer than translating a lexical sound.

3.2. Age group effects

The finding that older children responded more accurately and faster than their younger peers is consistent with many developmental studies showing improving behavioral efficacy with development (e.g., Zucchi and Bozzo, 1961; Trembach et al., 2004). What is perhaps more interesting is the fact that no electrophysiological differences were observed between the Pre-adolescent and Adolescent age groups at the level of N400 effect. In contrast, both the Pre-adolescent and Adolescent N400 effects were larger and later in comparison to Adults'[2].

---

[2] One could potentially attribute the observed developmental changes in the N400 effect amplitude to such non-functional phenomena such as skull size or bone density. However, by 7 years of age, a child's head is about 95% the size of an adult's head (e.g., Courchesne et al., 2000). These caveats are present in every single developmental study that compares young infants and children to adolescents and/or adults, and is not specific just to this study. As a result, the comparisons described here are free of systematic biases stemming from such issues.

This suggests that a significant developmental change in both verbal and non-verbal processing mechanisms occurs as late in development as late adolescence. This result both supports and contradicts previous developmental ERP findings.

Holcomb et al. (1992) reported N400 amplitude and latency reductions in children between the ages of 5 and 16 as they completed a word-by-word sentential semantic processing task. However, they did not report post-hoc measures specifying which age groups significantly differed. Thus, the narrower age range in the present study, not including the youngest (5 and 6 years) or the oldest (15 and 16 years) ages of the Holcomb et al. (1992) study, may have contributed to the study differences concerning age-related N400 effects.

However, Hahne et al. (2004) studied children of rather comparable ages as in the present study (7–8 and 10–13 years of age) and found that the N400 latency was longer in their younger than older groups during a continuous speech sentential semantic processing task, while we found no such difference. Both Holcomb et al. (1992) and Hahne et al. (2004) used auditory and/or visual sentences with either considerable time intervals between words (700–1000 ms; Holcomb et al., 1992) or as connected speech (Hahne et al., 2004). In these paradigms, children had to integrate multiple words of a sentence in order to decide whether the target word semantically fit or was anomalous. Therefore, it appears that the major determinant between the current and earlier results is the task differences leading to the engagement of distinct language-related processes.

Clearly, sentential designs involve more extended linguistic processing, which may both facilitate (richer context) and interfere (working memory load) with semantic integration. Therefore, those sentential paradigms may have enhanced developmental effects by compounding lexical semantic integration in its "pure", simplest form with more involved forms of cognitive and linguistic processing, such as working memory and syntactic knowledge. In contrast, in the present study, semantic context consisted of a single picture, largely overlapping in time with the following word or sound. Thus, memory load and grammatical skills were not taxed in this study. Based on the present results, it appears that in such a simple match/mismatch context, the neural response to a semantically anomalous stimulus is quite similar across the Pre-adolescent and Adolescent age groups. This may be due to the fact that this level of simplicity permitted even the youngest children in the study to perform as efficiently as their older peers, and is consistent with the finding that the N400 effect indexes an automatic stage of semantic processing (Cummings et al., 2006). Perhaps only when the task becomes more demanding, such as the processing of complex sentences, do the developmental semantic processing differences surface. Indeed, close inspection of raw data in the previous studies suggests that the N400 effect occurs later as the task demands become more complex (498 to 619 ms in Holcomb et al., 1992; 500 to 800 ms in Hahne et al., 2004).

Further evidence supporting the task-demand account of the N400 effect latencies comes from Byrne et al. (1999). Using a paradigm very similar to the present study's design, they presented children (5–12 years) with a picture that was either correctly or incorrectly aurally named. Within this experimental design, the N400 of 7- and 8-year-old children peaked at approximately 500 ms while the N400 of 11- and 12-year-old children peaked at approximately 400 ms, which is comparable to the N400 effect latencies recorded in the present study.

3.3. Sound type (Word/Environmental sound) × Age group interaction

At first glance, the fact that the Word N400 effect latency did not change with age while the Environmental Sound N400 effect latency gradually diminished is a counter-intuitive finding. This suggests that Environmental Sounds are not only processed faster than Words at all ages, but that processing efficiency of meaningful, non-verbal sounds improves throughout development. This claim is supported by behavioral results from a large-scale study involving 109 children (ages 6–18), in which the reaction times to environmental sound labels were somewhat longer and more variable in the younger children, but then showed a much steeper decline over age than did reaction times to word labels (Borovsky et al., submitted).

Above, we have already discussed the possibilities related to the overall longer processing time required for Words as compared with the Environmental Sounds. In fact, this experiment may have recruited optimally efficient word processing mechanisms, in part due to the multiple co-activations in the lexical domain exerted by the picture contexts resulting in an earlier developmental plateau. In contrast, processing of environmental sounds may lend itself to a greater degree of automaticity, and maximizing it might take longer developmental time. A related hypothesis would be that early in development, both verbal and meaningful non-verbal input are subjected to "checking in" with the lexical loop due to the dominance of the verbal processing devices during everyday functioning. It may be not until the high-order and high-efficiency automatic subroutines come online when such redundancy ceases or is significantly reduced.

The subtle time course differences in the processing of words and environmental sounds brings into question how different is semantic processing that is induced by a verbal label as compared to that induced by an environmental sound. The similarities suggest that processing both may rely on the same sensory, motor, and attentional precursors to language. If similar systems are involved in the processing of verbal and non-verbal meaningful sounds, the task described here could be useful in addressing the linguistic specificity of language impairment in neuro-developmental disorders, such as "specific language impairment", autism, or epileptic aphasia. If children from neuro-developmental populations showed diminished N400 effects to both words and environmental sounds, the argument could be made that these children have a generalized semantic processing impairment rather than one related to processing specifically of lexical material.

4. Conclusions

This study is the first to examine how Pre-adolescent and Adolescent children respond to two kinds of meaningful stimuli: Words and Environmental Sounds. Developmentally, the overall N400 effect decreased in magnitude from adolescence to

adulthood. This implies that significant maturational changes in the mechanisms underlying semantic processing occur through late adolescence. Further, the Word N400 effect latency remained constant across the age groups but that of the Environmental Sounds decreased from adolescence to adulthood. We here conjecture that this developmental change may be caused by the fact that verbal labeling may be an intermediate step in Environmental Sound identification until more efficient, more automatized strategies come online. This does not happen with Words because they are already well established by the age of 7 as the dominant, and automatic, device for semantic identification and interpretation.

5. Experimental procedure

5.1. Participants

Twenty-eight children in two age groups (Pre-adolescent: 7–9 years, $N = 13$, 6 male; Adolescent: 12–14 years, $N = 15$, 8 male) participated in the experiment. Eighteen of these children (8 male, 9 Pre-adolescent) completed the Verb Experiment and the other ten children (6 male, 4 Pre-adolescent) completed the Noun Experiment[3]. The data of fifteen Adults (8 male, 5 Noun) from the previous adult study (Cummings et al., 2006) were included in the Sound Type and Age Group comparisons. All participants were right-handed native speakers of American English. All children were screened and were found to have normal or corrected-to-normal hearing and vision. As a measure of general cognitive ability, all children were administered the Wechsler Intelligence Scale for Children-Third Edition (WISC-3). All children scored within or above the normal range as compared to other children their age (Mean Percentile Score = 73; SD = 23)[4]. All subjects signed informed consent in accordance with the UCSD Human Research Protections Program.

5.2. Study design

This study employed a picture-sound matching design to assess electrophysiological brain activity related to semantic integration as a function of auditory input type (Words vs. Environmental Sounds; Fig. 4). Colorful pictures of objects were presented with either a word or an environmental sound, with the sound starting at 600 ms after the picture onset and both stimuli ending together. Such a construct allowed for obtaining clean visual ERPs to the pictures, left no time for their conscious

labeling (verbalization), and avoided working memory load.[5] In half the trials, pictures and sounds matched, and in half the trials they were mismatched.

The design contained three levels of independent variables. The first independent variable was Sound Type (2 levels): Words and Environmental Sounds. A second independent variable was Trial Type: Match and Mismatch, in which every sound (either Word or Environmental) was once presented in conjunction with a picture that it correctly identified and once presented in conjunction with a mismatching picture. The third independent variable was Age Group: Pre-adolescent (7–9 years), Adolescent (12–14 years), and Adult (18–25 years) group. Subjects were randomly assigned to either the Noun or Verb Experiments, but the Environmental Sounds were the same in both Experiments.

5.3. Stimuli

Auditory stimuli were digitized at 44.1 kHz with a 16-bit sampling rate. The average intensity of all auditory stimuli was normalized to 65 dB SPL.

5.3.1. Environmental sounds
The sounds came from many different non-linguistic areas: animal cries ($n = 15$, e.g. cow mooing), human non-verbal vocalizations ($n = 3$, e.g. sneezing), machine noises ($n = 8$, e.g. car honking), alarms/alerts ($n = 7$, e.g. phone ringing), water sounds ($n = 1$, toilet flushing), event sounds ($n = 12$, e.g. bubbles bubbling) and music ($n = 9$, e.g. piano playing). The sounds ranged in duration from 400–870 ms (mean = 574 ms, SD = 104 ms).

5.3.2. Words
In an attempt to make the word stimuli comparable to the environmental sounds, words were pronounced by three North American speakers (one female and two male), which allowed for a greater acoustic variation. The words were digitally recorded in a sound isolated room (Industrial Acoustics Company, Inc., Winchester, UK) using a Beyer Dynamic (Heilbronn, Germany) Soundstar MK II unidirectional dynamic microphone and Behringer (Willich, Germany) Eurorack MX602A mixer. Noun stimuli ranged in duration from 262–940 ms (M = 466 ms, SD = 136 ms) and the verb stimuli were 395–1154 ms (M = 567 ms, SD = 158 ms).

5.3.3. Visual stimuli
Pictures were full-color, digitized photos (280 × 320 pixels) of common action-related objects that could produce an environmental sound and be described by a verb or a noun. All pictures extended 6.9° visual angle and were presented in the middle of the computer monitor on a gray background. The same set of pictures was used for both the matching and

---

[3] The study was initially designed to address children's linguistic processing abilities. It was assumed that verbs would be more appropriate because they are better correlates of environmental sounds (i.e., associated with actions) and are acquired later than nouns. However, our adult data (Cummings et al., 2006) did not show a difference between noun and verb N400 effect. Thus, we recruited a minimal number of children examine the word class effect, based on the notion that differences that are not observed in adults could potentially be seen in children.

[4] The children's raw scores on the Verbal Vocabulary (Mean = 13; SD = 2.3) and Verbal Similarities (Mean = 13.4; SD = 3.1) subtests of the WISC-3 were both within normal limits.

[5] The 600 ms window was based on evidence by Simon-Cereijido et al. (2006), who reported that the time it took for an adult to identify and process a familiar object and then formulate a verbal label for that object was approximately 800–1000 ms. Additionally, Dick et al. (2002) reported reaction times for when subjects were asked to subvocally label environmental sounds (approximate range: 850–2000 ms) and also when subjects were not asked to label the sounds (approximate range: 750–2000 ms), both of which far exceeded 600 ms window between picture onset and sound onset.

Noun Word Trials

Verb Word Trials

Environmental Sound Trials
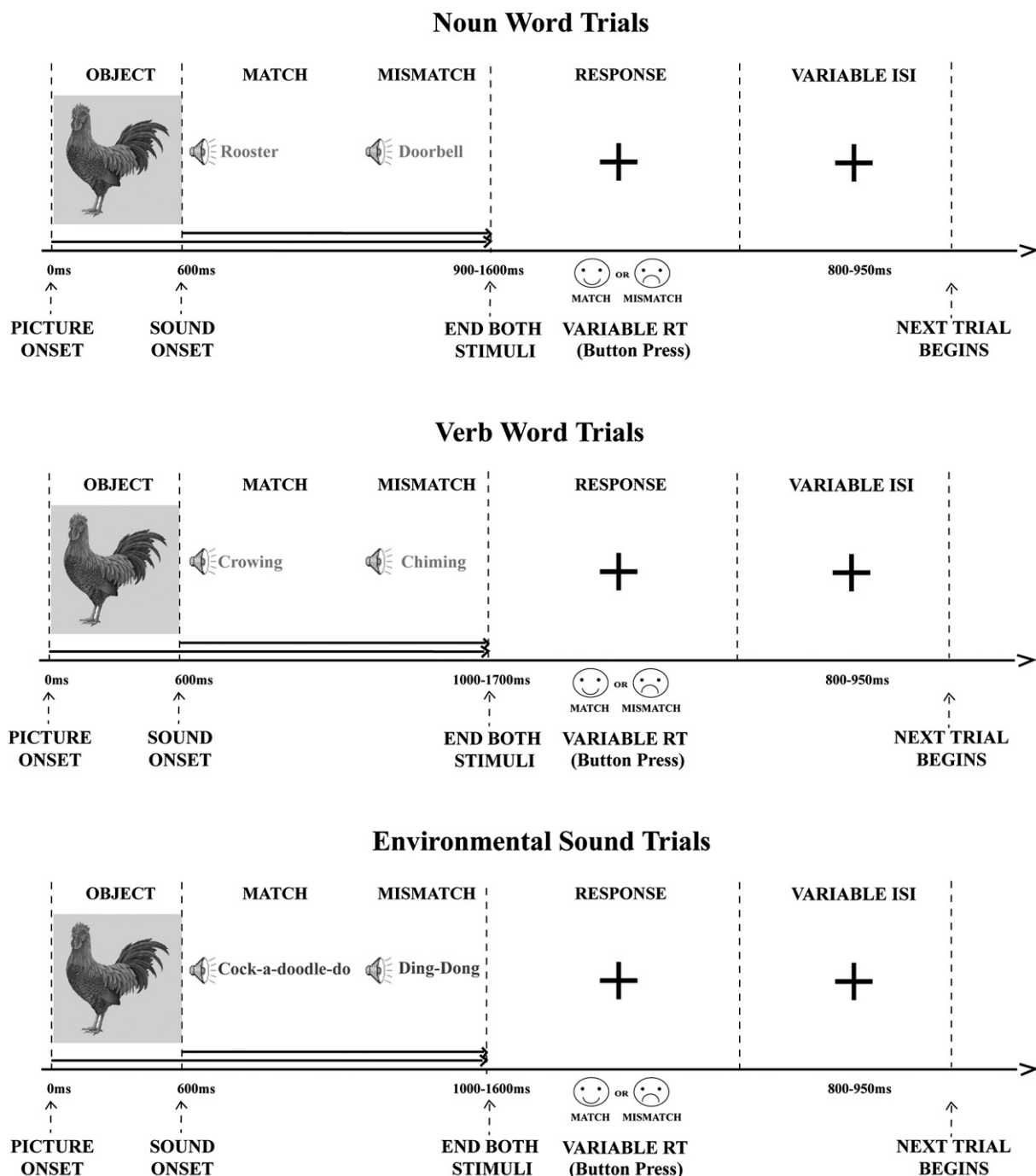
Fig. 4 – Experimental design. A picture of a real object was presented on a computer screen for 600 ms. The picture remained visible while a word or environmental sound was presented via loudspeakers. The auditory stimulus either matched or mismatched the visual stimulus. Once the auditory stimulus was complete, the picture disappeared from the computer screen and was replaced by a fixation cross, which remained visible until subjects recorded their response with a button press. Trial length was variable, being a combination of stimulus presentation (approximately 900–1700 ms), reaction time ($M = 790$–1079 ms), and variable inter-stimulus interval (ISI; 800–950 ms), which was calculated from the onset of the subject's button press. The happy face sticker always represented a match trial response and the sad face sticker always represented a mismatch trial response. Half of the subjects had the happy face button on the right side of the button box, and the other half had it on the left side.

mismatching picture contexts. The only constraint on the semantic mismatching trials was that the mismatch had to be unambiguous (e.g. the picture of a basketball was NOT presented with the sound of hitting a golf ball).

For the sake of between-groups comparison, this experiment used the same paradigm as an earlier adult experiment (Cummings et al., 2006), which included control trials of non-meaningful stimuli. The non-meaningful sounds were matched

in duration to the environmental sounds and consisted of computer-generated sounds that were not easily associated with any concrete semantic concept. In order to permit picture–sound matching, the non-meaningful sounds were chosen to portray either a 'smooth' sound, e.g. a harmonic tone, or a 'jagged' sound, e.g. cracking noise. However, the non-meaningfulness dimension of the task was not a focus of the present experiment and these trials were not analyzed for the present paper.

5.4. Procedure

The stimuli were presented in blocks of 54 trials. All trial types occurred with equal probability, in a pseudo-random order with the constraints that no picture was presented twice in a row and that certain sounds could not be mismatched with certain pictures. Thus, there would have been no immediate stimulus priming across word and environmental sound trials, or across matching and mismatching trials. Six stimulus blocks (324 trials altogether) were presented to every subject. The pictures and sounds were delivered by stimulus presentation software (Presentation software, Version 0.70, www.neurobs.com). Pictures were presented on a computer screen situated 120 cm in front of the subject and sounds were played via two loudspeakers situated 30° to the right and left from the midline, also in front of a participant. The sounds were heard as appearing from the midline space. The subjects' task was to press a button marked by the smiley face as quickly as possible if they thought the picture and auditory stimulus matched, and to press a button marked by the sad face if they thought that the stimuli mismatched. Response hands were counterbalanced across the subjects.

5.5. Behavioral data analysis

Accuracy on a given trial was determined by a subject's button press response to the Picture–Sound Type pair. Reaction time was computer-calculated from the onset of the auditory stimulus to the computer's registering of the button press. Only reaction times from correct trials were included in the analyses. Since there was a wide variation in reaction times, median reaction times were computed within subjects for each condition and used in data analysis (means yielded comparable results). Responses of subjects from both Word Class experiments (Noun/Verb) were pooled for the Sound Type and Age Group analyses. The effects of Age Group on Sound Type were examined in an Age Group (Pre-adolescent/Adolescent/Adult)×Sound Type (Words/Environmental Sounds)×Trial Type (Match/Mismatch) ANOVA.

5.6. EEG recording and averaging

Continuous EEG was recorded using a 32-electrode cap (Electrocap, Inc.) with the following electrodes attached to the scalp, according to the International 10–20 system: FP1, FP2, F7, F8, FC6, FC5, F3, Fz, F4, TC3, TC4 (30% distance from T3-C3 and T4-C4, respectively), FC1, FC2, C3, Cz, C4, PT3, PT4 (halfway between P3-T3 and P4-T4, respectively), T5, T6, CP1, CP2, P3, Pz, P4, PO3, PO4, O1, O2, and right mastoid. Eye movements were monitored with two electrodes, one attached below the left eye and another at the corner of the right eye. During data

acquisition, all channels were referenced to the left mastoid; offline, data was re-referenced to the average of the left- and right-mastoid tracings.

The EEG (0.01–100 Hz) was amplified 20,000×and digitized at 250 Hz for the offline analyses. Prior to averaging, an independent-component analysis (ICA; Jung et al., 2000) was used to correct for eye blinks and lateral eye movements. The remaining artifactual trials due to excessive muscle artifact, amplifier blocking, and overall body movements were rejected from further analyses. Epochs containing 100 ms pre-auditory stimulus and 900 ms post-auditory stimulus time were baseline-corrected with respect to the pre-auditory stimulus interval and averaged by stimulus type: Picture–Word Match, Picture–Word Mismatch, Picture–Environmental Sound Match, and Picture–Environmental Sound Mismatch. A low-pass Gaussian digital filter was used to filter out frequencies higher than 60 Hz. On average, the remaining individual data contained 95 (SD=10) Word trials (Pre-Adolescent $M=86$ (SD=11); Adolescent $M=95$ (SD=8); Adult $M=101$ (SD=3)) and 98 (SD=9) Environmental Sound trials (Pre-Adolescent $M=87$ (SD=9); Adolescent $M=98$ (SD=8); Adult $M=103$ (SD=4)). The N400 effect was measured using difference waves for each sound type, which were made by subtracting the ERP responses of the matching trials from the ERP responses of the mismatching trials (e.g. Picture–Word Mismatch – Picture–Word Match; Fig. 1, right column)[6].

5.7. ERP measurements

Only ERP responses from correct trials were included in the analyses. The amplitude and latency of the N400 effect were measured from the difference waves time-locked to the onset of the auditory stimuli. The N400 effect was measured from 300–500 ms at the 23 electrodes, where it was present in the grand-average waveforms: F3/F4, Fz, FC5/FC6, FC1/FC2, TC3/TC4, C3/C4, Cz, PT3/PT4, CP1/CP2, T5/T6, P3/P4, Pz, and PO3/PO4. First, peak latencies of the N400 effect were measured from the grand averages of the three Age Groups. A "center" latency of the N400 effect for each group was then calculated as a mean of the latencies of the 23 electrodes. Mean amplitudes of the N400 effect of each subject were measured at each of the 23 electrodes from 25 ms before to 25 ms after the "center" latency of the group to which the subject belonged. Thus, the N400 effect amplitudes for each subject represented means over 50 ms at the 23 electrodes and the N400 effect latencies represented peak latencies of the measured peaks (e.g., Federmeier et al., 2007).

Two-tailed independent-sample t-tests were conducted to test the significance of the N400 effect elicited by the two sound types at the midline electrodes. Mixed (within- and between-group) ANOVA analyses were performed using all 23 measured electrodes. The effects of Age Group on Sound Type were examined in an Age Group (Pre-adolescent/Adolescent/Adult)×

---

[6] The N400 is an ERP component that is present in both matching and mismatching ERP waveforms (Figs. 1 and 2 – left two columns). The N400 effect is the result of subtracting the matching ERP waveform from the mismatching ERP waveform. By subtracting out the common responses to both types of stimuli, only the difference between them remains, thus making the N400 effect very pronounced (Fig. 1 – right column).

Sound Type (Words/Environmental Sounds) × Electrode (23 levels) ANOVA. When applicable, Geiser-Greenhouse corrected p-values are reported.

Scalp distribution analyses for Age Group and Sound Type effects were completed to search for differences along the Laterality and Anterior–Posterior dimensions. For the Anterior–Posterior analyses, mean amplitudes from 14 electrodes comprising 7 anterio-posterior levels of scalp distribution were included as follows: F3/F4, FC1/FC2, C3/C4, CP1/CP2, P3/P4, PO3/PO4, and O1/O2. The Laterality analyses utilized mean amplitudes from 8 electrodes comprising 2 levels of left-right scalp distribution: FC5/FC6, TC3/TC4, PT3/PT4, and T5/T6. The mean amplitudes from these 22 electrodes were normalized using a z-score technique calculated separately for each Age Group's responses to Words and Environmental Sounds (Picton et al., 2000). All scalp distribution analyses involving interactions between variables were completed using normalized amplitudes. Normalized amplitudes were used only for between-group scalp distribution analyses.

For both the Anterior-Posterior and Laterality analyses, the scalp distribution differences were examined between the Age Groups as a function of Sound Type: Age Group (Preadolescent/Adolescent/Adult) × Sound Type (Word/Environmental Sound) × Anteriority–Posteriority (7 levels) or Laterality (2 levels).

Acknowledgments

REFERENCES

Ballas, J.A., 1993. Common factors in the identification of an assortment of brief everyday sounds. J. Exp. Psychol. Hum. Percept. Perform. 19 (2), 250–267.

Ballas, J.A., Howard, J., 1987. Interpreting the language of environmental sounds. Environ. Behav. 19 (1), 91–114.

Byrne, J., Connolly, J., MacLean, S., Dooley, J., Gordon, K., Beattie, T., 1999. Brain activity and language assessment using event-related potentials: development of a clinical protocol. Dev. Med. Child Neurol. 41, 740–747.

Courchesne, E., Chisum, H., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hinds, S., Press, G., 2000. Normal brain development and aging: quantitative analysis at in Vivo MR Imaging in healthy volunteers. Radiology 216, 672–682.

Cummings, A., Čeponienė, R., Koyama, A., Saygin, A.P., Townsend, J., Dick, F., 2006. Auditory semantic networks for words and natural sounds. Brain Res. 1115, 92–107.

Cycowicz, Y.M., Friedman, D., 1998. Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. Brain Cogn. 36 (1), 30–51.

Dick, F., Bussiere, J., Saygin, A.P., 2002. The effects of linguistic mediation on the identification of environmental sounds. Cent. Res. Lang. Newsl. 14 (3).

Dick, F., Saygin, A.P., Moineau, S., Aydelott, J., Bates, E., 2004. Language in an embodied brain: the role of animal models. Cortex 40 (1), 226–227.

Federmeier, K., Wlotko, E., de Ochoa-Dewald, E., Kutas, M., 2007. Multiple effects of sentential constraint on word processing. Brain Res. 1146, 75–84.

Gygi, B., 2001. Factors in the identification of environmental sounds. Unpublished Doctoral Dissertation. Department of Psychology and Cognitive Science. Indiana University, Bloomington. PhD: 187.

Hahne, A., Eckstein, K., Friederici, A., 2004. Brain signatures of syntactic and semantic processes during children's language development. J. Cogn. Neurosci. 16 (7), 1302–1318.

Holcomb, P., Coffey, S., Neville, H., 1992. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. Dev. Neuropsychol. 8 (2&3), 203–241.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111 (10), 1745–1758.

Juottonen, K., Revonsuo, A., Lang, H., 1996. Dissimilar age influences on two ERP waveforms (LPS and N400) reflecting semantic context effect. Cogn. Brain Res. 4 (2), 99–107.

Kutas, M., Hillyard, S.A., 1980a. Event-related brain potentials to semantically inappropriate and surprisingly large words. Biol. Psychol. 11 (2), 99–116.

Kutas, M., Hillyard, S.A., 1980b. Reading between the lines: event-related brain potentials during natural sentence processing. Brain Lang. 11 (2), 354–373.

Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. 11 (5), 539–550.

MacWhinney, B. (Ed.), 1999. The emergence of language. Lawrence Erlbaum Associates, Mahwah, NJ.

Neville, H., Kutas, M., Chesney, G., Schmidt, A., 1986. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. J. Mem. Lang. 25, 75–92.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37 (2), 127–152.

Pinker, S., 1994. The language instinct. William Morrow & Co., New York, NY.

Plante, E., van Petten, C., Senkfor, A.J., 2000. Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. Neuropsychologia 38 (13), 1669–1684.

Saygin, A.P., Dick, F., Wilson, S.M., Dronkers, N.F., Bates, E., 2003. Neural resources for processing language and environmental sounds: evidence from aphasia. Brain 126 (Pt 4), 928–945.

Saygin, A.P., Dick, F., Bates, E., 2005. An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. Behav. Res. Methods 37 (1), 99–110.

Shafiro, V., Gygi, B., 2004. How to select stimuli for environmental sound research and where to find them. Behavior Research Methods, Instruments, & Computers. Special Issue: Web-based archive of norms, stimuli, and data: Part 2, 36(4), pp. 590–598.

Simon-Cereijido, G., Bates, E., Wulfeck, B., Cummings, A., Townsend, J., Williams, C., Čeponienė, R., 2006. Picture naming in children with Specific Language Impairment: Differences in neural patterns throughout development. Poster presented at

the Annual Symposium on Research in Child Language Disorders. Madison, WI.

Thelen, E., Smith, L., 1994. A dynamic systems approach to the development of cognition and action. MIT Press, Cambridge, MA.

Trembach, A., Belyaev, M., Lysenko, V., 2004. Age-related changes in attention and impulsivity in young schoolchildren. Hum. Physiol. 30 (5), 537–533.

Van Petten, C., Rheinfelder, H., 1995. Conceptual relationships between spoken words and environmental sounds: event-related brain potential measures. Neuropsychologia 33 (4), 485–508.

Zucchi, M., Bozzo, M., 1961. The measurement of simple reaction time at the age of development with apparatus with automatic presentation of stimuli (translated). Arch. Psicol. Neurol. Psichiatr. 22, 547–557.