NEUROLINGUISTIC RESPONSES TO PERCEPTION OF SPEECH IN INCONGRUENT PICTURE CONTEXT

Eeva Klintfors, Ellen Marklund & Francisco Lacerda

Department of Linguistics, Section for Phonetics, Stockholm University, Sweden
eevak@ling.su.se; ellen@ling.su.se; frasse@ling.su.se

ABSTRACT

This study investigates N400, the event related potential (ERP) that reflects semantic processing in cortex. The stimuli were semantically congruent *vs.* incongruent audio-visual pairings of objects (*i.e.* words presented in familiar picture context). As predicted, the results showed in time extended significant peak-to-peak differences between congruent and incongruent picture-word pairings at the centroparietal and the parietal recording sites. The rationale for this study is our pioneer intent to verify use of ERPs in response to the current materials. After extension of this research, the data will be used in comparison to answer questions on appearance of the semantic component in young children. Based on earlier research, we expect that the N400 component in children will be greater in amplitude, delayed in latency and more widely distributed in scalp distribution.

Keywords: neuroimaging, semantic incongruity

1. INTRODUCTION

This study is part of the "Early Development of Hemispheric Specialization for Speech Processing" project. The general aim of the project is to study establishment of language-related specialization in the brain and its specific links to different phases of language development early in life. To achieve this goal a combination of neurophysiological EEG (electroencephalography) and behavioral (eye-tracking) data to assess spatial and temporal structure of the subjects' cortical activity while exposed to different linguistics tasks will be collected. Another immediate goal of this project is to apply the results from the basic research to establish potential early signs of deviant language development in children with autism spectrum disorders (ASD) and to propose valid linguistic strategies that may promote the development of functional language capabilities among them.

As a first step towards these goals, perceptual experiments with adults will be performed to investigate neurological correlates to semantic representation. Collection of data from competent language users makes it possible to establish baselines against which descriptions of early typical (normal), as well as identification of deviant language development may be tested.

This study provides data on cortical networks that are likely to mediate semantic processing as revealed by the so called semantic incongruity effect. More specifically, adult responses to semantic incongruities, presented in the auditory (A) and visual (V) sensory modalities was tested in two situations: one using congruent, and the other incongruent A-V pairings of objects.

2. BACKGROUND

2.1. AV-processing in adults and children

It is common knowledge that information in the environment is taken in through our senses and that combined input from sensory channels contribute to the contents of concepts. However, information processing by a child is substantially different relative to a linguistically experienced system. For example:

(1) Perception of speech sounds in infancy is general in the sense that speech sounds are perceived as sounds *per se*, relative to a linguistic categorization of sounds in adults. That is, infants are able to discriminate all the phonetic contrasts they have been tested on [14, 20], while this capacity in adults applies to native contrasts only [22-24].

(2) Speech perception in adults is typically facilitated by seeing the speaker's articulation of speech [4, 6, 10, 21] while it is not clear when corresponding ability is present in children. In a classic study, it was demonstrated that 4 to 4.5 month old infants were able to identify articulatory movements for a vowel out of two alternatives (/a/ and /i/) [15-17]. However, an extension of these studies, using a design with four alternatives (/ba/, /by, /a/, /y/) that reduces the chance level from 50 to 25%, could not verify matching of audio-articulation in 6 to 8 month-old infants [19].

(3) Emergence of linguistically induced semantic representations are known to appear sometime during the second year of life [2], but the mechanisms involved in integration of meaning at a higher level are not yet well understood in children. Corresponding imaging studies that aim to identify cortex regions sensitive to semantic variations are discussed in the next section.

2.2. Semantic ERPs in adults and children

N400 is a well established ERP component known to be sensitive to variations in the semantic content of words and sentences [3, 11, 18]. As indicated by the name of the component, it has a negative peak at around 400ms post-stimulus. For example, all words presented along visual images induce an N400 effect, but the amplitude of it is highest when the word is semantically incongruent with the image.

An extensive study (N=130, 5 to 26 years) on age dependent appearance of N400 showed that both children and adults produce a larger N400 in response to semantically anomalous relative to well-formed sentences [12]. The authors concluded that N400 activation in younger children (5 to 14 years) is widely distributed, including frontal electrodes, relative to activation in adults and older children where it is limited to posterior areas only. Other studies, reporting significantly larger N400 amplitude components for children (5 to 13 years) relative to adults, have also found significant effects for delayed N400 peak latency and extended duration of N400 in children [1, 7, 13].

The few studies performed on the presence of N400 in infants suggest that semantic processing mechanisms indexed by N400 are present at 14 and 19 months of age [8, 9].

2.3. The present study

The present study investigates an adult N400-response using a child-adapted experiment design, creating a baseline for future studies. A picture-word paradigm with coloured pictures of familiar objects and slowly spoken words, either naming the object in the picture (A-V congruent condition) or naming another object (A-V incongruent condition) is used. The latter condition is expected to elicit the most significant N400-responses (see Figure 1). The distribution of N400 is expected to be maximal over posterior midline scalp locations, in specific the centroparietal (CPz) and parietal (Pz) sites.

3. METHOD

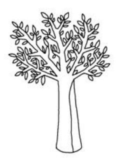
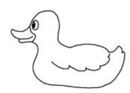
3.1. Participants

The participant was faculty staff of the Linguistic Department, Stockholm University (N=1; female, age 35 years). She was unfamiliar with the materials and the research question of the study. She was native speaker of Swedish, right-handed, and had normal hearing. The participant did not receive anything in exchange for her participation.

3.2. Stimuli and experimental design

Stimuli consisted of images of objects and words naming the objects, paired either congruently or incongruently. Incongruent pairs were assigned to control for voicing and articulation type of the first phoneme. The words were chosen from the Swedish Early Communicative Developmental Inventory (SECDI) [5] for compatibility with future infant studies, in which age-appropriate words will be chosen. Since adults can be assumed to understand all words in the inventory questionnaire, 48 words naming different objects were selected arbitrarily.

Figure 1: The test included two types of AV-pairings: incongruent word in picture context, illustrated in the upper part of the figure *vs.* congruent word in picture context, illustrated in the lower part of the figure. A larger N400 effect was expected in response to the upper incongruent A-V pairing.

Picture	Audio
	”duck”
	”duck”

The words were read by a female native speaker of Swedish and recorded in an anechoic chamber, using a ½-inch Brüel & Kjær free-field condenser microphone (4190) and a Brüel & Kjær amplifier (2209). The materials were digitally saved to a hard drive (Creative Audigy 2ZS sound card) with a sampling rate of 44100Hz. The words were segmented and manually matched for loudness using Cool Edit 2000. Single word duration ranged from 526 to 1178 ms. The pictures were drawings of the selected objects centred on a grey background.

The experiment was designed in E-Prime 2.0. Each picture was introduced for 1000 ms before the word was presented, and remained on the screen for a total of 5000 ms. Each picture was used twice, once in the semantically congruent condition and once in the incongruent condition, resulting in a total of 96 randomised trials. The test took approximately 14 minutes.

3.3. Procedure

To measure stimulus-related brain activation, EEG-head-net (Electrical Geodesics Inc.) with 128 sensor channels was used, size selected according to head circumference. The midline frontal (Fz), frontal-central (FCz), centroparietal (CPz), and parietal (Pz) sites were included in the analyses. Each scalp side was referred to the linked mastoids. All electrodes' impedance was lower than 5 kΩ. Electrodes were placed above and below the left eye and at the outer canti to monitor blinks and eye movements.

After installation of the EEG head cap, the participant was comfortably seated on a padded chair in a dimly lit sound-attenuated room. She was advised to reduce eye blinking as much as possible. To avoid fatigue towards end of the test, the participant was asked to attentively listen to the audio and look at the pictures shown on the screen. The stimuli were presented via a pair of external computer speakers and the sound level of stimuli was approximately 60 to 80 dB. The experiment was run using E-Prime 2.0 and Net Station 4.2.

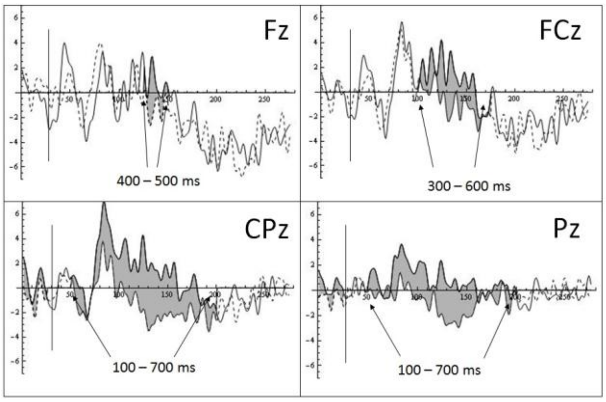
3.4. Data preparation and analysis

Net Station tools cleaned the EEG data with a band pass filter from 0.3–30 Hz noise, segmented the stimulus responses into 96 segments of 1100 ms each (100 baseline before stimuli onset, 1000 ms after stimuli onset), and removed unusable channels and segments (*e.g.* eye blinks) before collating segment averages across stimuli conditions, and referencing the EEG-voltage measurements to a baseline prior to stimuli onset.

4. RESULTS

Figure 2 displays ERPs elicited by congruent (solid lines) *vs.* incongruent (dotted lines) A-V pairings for the time period beginning 100 ms prior to the stimulus and ending 1000 ms post-stimulus. Waveforms for four electrode sites are depicted (Fz, FCz, CPz, and Pz) to highlight the scalp distributions for the N400 waveforms.

Figure 2: Averaged ERP waveforms generated by the congruent (solid lines) *vs.* incongruent (dotted lines) A-V pairings at four electrode sites. Shaded areas and to them corresponding arrows represent time periods during which the two waveforms are significantly diverging. Stimulus onset (vertical lines) is shown corresponding to 0 ms.



As visual inspection suggested that the incongruity effect was larger at the centroparietal (CPz) and parietal (Pz) sites relative to the midline frontal (Fz) and frontal-central (FCz) sites, we performed two-way analyses for each of them. For the frontal sites the incongruity effect reached significance only from 400 to 500 ms at the Fz, and from 300 to 600 ms at the FCz, for the both parietal sites, at CPz and Pz, it was present within 100 to 700 ms.

5. DISCUSSION

As in previous studies with adults, our results showed a significant N400 effect in response to semantic incongruities that was maximal over posterior midline scalp locations. Due to the pilot character of the study, it is premature to speculate on implications of these results given the obvious need for replication and extension of the current research. Also, additional analyses of cortical responses *within* the incongruent condition are to be performed on animate *vs.* inanimate objects. However, if these results prove robust then they could be very informative and function as a baseline for N400 studies in children with typical language development *vs.* in children with ASD.

The current results for the N400 component are very similar to previous findings in adults. Yet, in line with earlier studies with young children, the dominant differences between adults and infants are expected to be in N400 component amplitude, more delayed component latency, and more widely distributed scalp distribution relative to adults.

To be tested in a longitudinal design, differences in cortical activation patterns between semantically congruent *vs.* incongruent A-V materials are hypothesized to be age-dependent. A few investigators have demonstrated the presence of N400 at 14 and 19 months, yet more studies are needed to explore the onset of the effect in infants.

The onset of speech and other linguistic milestones are typically delayed in children with ASD. Autism is of particular interest in this project because it offers the opportunity of testing a theoretical perspective proposing that correlated sensory information is the very key to the development of linguistic referential function. The activation pattern found in typically developing children is not expected to show as clearly (if at all) in children with ASD.

6. ACKNOWLEDGEMENTS

Research supported by The Swedish Research Council (nr 2009-2245).

7. REFERENCES

[1] Atchley, R.A., et al. 2006. A comparison of semantic and syntactic event related potentials generated by children and adults. *Brain and Language* 99, 236-246.

[2] Bates, E., et al. 1994. Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* 21, 85-123.

[3] Bentin, S., McCarthy, G. 1994. The effects of immediate stimulus repetition on reaction times and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology, Learning Memory, and Cognition* 20, 130-149.

[4] Dekle, D.J., Fowler, C., Funnel, M.G. 1992. Audiovisual integration in perception of real words. *Perception and Psychophysics* 51, 355-361.

[5] Eriksson, M., Berglund, E. 1996. Swedish Early Communicative Development Inventory - words and gestures. *First Language* 19, 55-90.

[6] Fowler, C. 1996. Listeners do hear sounds, not tongues. *Journal of Acoustical Society of America* 99, 1730-1741.

[7] Friederici, A.D., Hahne, A. 2004. Development patterns of brain activity reflecting semantic and syntactic processes, In Weissenborn, J., Houle, B. (eds.), *Approaches to Bootstrapping: Phonological, Lexical, Syntactic, and Neurophysiological Aspects of Early Language Acquisition.* Amsterdam/Philadelphia: John Benjamin, 231-246.

[8] Friedrich, M., Friederici, A.D. 2004. N400-like Semantic Incongruity Effect in 19-Month-Olds: Processing Known Words in Picture Contexts. *Journal of Cognitive Neuroscience* 16(8), 1465-1477.

[9] Friedrich, M., Friederici, A.D. 2005. Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport* 16(6), 653-656.

[10] Green, K., Kuhl, P. 1991. Integral processing of visual place and auditory voicing information during phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* 17, 278-288.

[11] Hagoort, P., Brown, C.M. 2002. ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia* 38, 1518-1530.

[12] Holcomb, P.J., Coffey, S.A., Neville, H.J. 1992. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology* 8, 203-241.

[13] Juottonen, K., Revonsuo, A., Lang, H. 1996. Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantics context effect. *Cognitive Brain Research* 4, 99-107.

[14] Kuhl, P. 2004. Early language acquisition: Cracking the speech code. *Nature Reviews: Neuroscience* 5(11), 831-843.

[15] Kuhl, P., Meltzoff, A.N. 1982. The bimodal perception of speech in infancy. *Science* 218, 1138-1141.

[16] Kuhl, P., Meltzoff, A.N. 1984. The intermodal representation of speech in infants. *Infant Behavior and Development* 7, 361-381.

[17] Kuhl, P., Williams, K., Meltzoff, A.N. 1991. Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance* 17, 829-840.

[18] Kutas, M. 1997. Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology* 34, 383-398.

[19] Lacerda, F., et al. 2005. Emerging linguistics functions in early infancy. *The Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Nara, Japan: Lund University Cognitive Studies.

[20] Polka, L., Werker, J. 1994. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20(2), 421-435.

[21] Remez, R.E., et al. 1998. Multimodal perceptual organization of speech: Evidence from tone analogs of spoken utterances. *Speech Communication* 26, 65-73.

[22] Tees, R.C., Werker, J. 1984. Perceptual flexibility: Maintenance of recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology* 38(4), 579-590.

[23] Werker, J., Tees, R.C. 1999. Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology* 50, 509-535.

[24] Werker, J., Tees, R.C. 2002. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 25, 121-133.