Prosodic processing in the first year of life: an ERP study

Linda Garami[1], Anett Ragó[2,1], Ferenc Honbolygó[1,2], Valéria Csépe[1]

[1] Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary
[2] Institute of Psychology, University of Eötvös Loránd, Budapest, Hungary
garami.linda@ttk.mta.hu, www.humlab.cogpsyphy.hu

Abstract

From early months of life prosody has a prominent contribution to segmentation: prosodic boundaries overlap with syntactic ones and facilitate the extraction of syntactic regularities both at word and at phrase level. Therefore, the long-term representation of rhythmic features of the native language, especially the stress templates derived from regularities are assumed to play a particular role in pre-lexical processing. We examined the nature of early stress representation in a language with a fixed stress pattern in an electrophysiological experiment (acoustic passive odd-ball paradigm, 10 month-olds: 28 infants; 6 month-olds: 21 infants, 400 items, deviant: p=20%) using bi-syllabic Hungarian pseudo-words to follow how prosodic features contribute to processing saliency and how word stress templates based on regularities may emerge. We used legally and illegally stressed stimulus both in standard and deviant positions in separate conditions.

In the legal standard condition two mismatch responses (MMRs) temporally synchronized to each syllable could be recorded. On the contrary, in the illegal standard condition no significant response was found. It seems that language environment influences the processing of speech prosody and the MMR correlates of word stress processing are related both to saliency and to stress templates emerging during the first year of life.

Index Terms: language development, prosody, mismatch response

1. Introduction

Infants acquire their native language quickly and accurately among suboptimal conditions, as the utterances of adults around them are usually imperfect, non-segmented and changing with personal characteristics. One of the first problems infants need to solve is how to extract meaningful units from the continuous speech stream, without having any lexical knowledge. Originally Gleitman and Wanner [1] then later Anne Christophe and her colleagues [2] proposed that since prosodic cues like lexical stress may reliably signal word boundaries they can act as bootstrapping mechanisms by helping to locate words in the speech stream. These presumed prosodic bootstrapping mechanisms help infants to generate and apply rules extracted from distributional patterns of spoken language [3, 4, 5]. Therefore prosody, among other cues, such as allophonic, phonotactic and statistical or distributional cues [6], is assumed to facilitate early language acquisition. The extraction of prosodic regularities is based on acquiring the rhythmic segmentation procedure that allows infants segmenting their first patterns [7, 8, 9] which soon issue in that infants show a strong trochaic bias in languages that apply this pattern (i.e. English, German or Hungarian) in contrast to infants whose native language applies an iambic one. This implies that infants are able to detect specific patterns of distinct acoustic features related to prosody, and that the exact nature of this mechanism is mostly dependent on the native language specificity. Our research focuses on the development of representation resulted from extracting language-specific patterns.

The early segmentation hypothesis relies on adults' segmenting ability. According to the electrophysiological data of Honbolygó and Csépe the perception of stress pattern is based on long-term, pre- lexical, language-specific representations of stress information among Hungarian adults that they called stress templates [10]. Development of such a template should occur already in early infancy, as infants are sensitive to specific salient acoustic features and are able to extract rules given in patterns quite early. However, neither the relationship of these acoustic cues nor the developmental timing of them is clear yet. In our study we were interested in when the representation of native language's specific prosodic pattern signaling word boundaries could be abstracted and how these patterns are utilized when listening to new utterances. Hungarian language has some special characteristics as it is a syllable-based language like French, although has a strong stress initial rule as German, while vowel duration is a segmental feature in contrast to both.

Studies using electrophysiological methods refined our understanding of prosodic information by examining the underlying mechanisms without relying on behavioural responses. This method enables us to filter out the behavioral readiness of infants, as it measures only the brain activity correlating with automatic linguistic process that is independent of willpower. Most of the electrophysiological experiments applying the method of passive oddball paradigm are looking for deflections in event-related brain potentials (ERP) related to stress processing. Among adults, the Mismatch Negativity (MMN) event-related brain potential component is elicited when an unexpected change occurs in the auditory environment. The expectations can be generated either via forming a short-term memory trace of the actual stimulus set (with a series of identical stimuli) or resulted from this trace affected by long-term memory representations [11]. The unexpected change of suprasegmental information (prosody) evokes mismatch responses, where latency, amplitude, and polarity, all depend on the familiarity of the stimulus [12]. A cross-linguistic study [13] underpinned that language environment promotes different discrimination abilities no later than at the age of 4 months as French and German infants showed different ERP waveforms to different stress patterns by four-months of age. This bias was congruent with the deviation of language related stress cues and regularities: while in German the dominant stress pattern is stress on the first syllable, in French stress is on the second one (among bi-syllabic words). In an oddball paradigm infants responded with a positive mismatch response (MMR) only to the illegally stressed pseudo-words of their own native

10.21437/SpeechProsody.2014-207

language. The authors used salient cues typical for German such as duration and intensity. However, syllabic stress is not determined by the same acoustic features in other languages, so stress is not uniform from an acoustic point of view. For example infants show high sensitivity to duration [14, 15] very early, while intensity or other features may show a different developmental trajectory. Dividing salient acoustic feature processing and prosodic template processing is a critical issue in all languages. The most important acoustic features defining stress pattern vary across languages, so discriminating phoneme duration or pitch or f0 alone does not allow us to shed light on detecting stress pattern deviations. This ability is derived as a rule extraction based on the early language experience. The ERP studies focusing on prosodic information varied mostly the duration of speech sounds; consonants [16] or vowels [13, 17, 18]. In Hungarian vowel duration is a segmental cue, where young learners have to extract a typical stress pattern while taking complex variations (acoustic perturbations in case of long vowels) into account. However, we took advantage of the expressed regularity characteristic for Hungarian with word-initial stress for all words (except for compound words), a pattern correlating with linguistic units, and therefore providing a strong unvarying environment for forming stress-related expectations [19, 20].

As lexical stress used in the Hungarian language (always word-initial) is associated perfectly with the words' beginning, this allows us to follow how these features contribute to processing saliency and how word stress templates based on regularities might emerge.

The objective of our ERP study was to reveal the ability 6 and 10 month-old infants to use acoustically rich stress information, including pitch and f0 changes, the two most important features characteristic for syllabic stress in Hungarian. We assumed that besides processing salient distinct features only, 6 month old infants have already extracted a specific rule regarding the stress pattern of their native language and violation is detected on a higher level as a long- term representation of stress.

2. Methods

We used the experimental paradigm of Honbolygó and Csépe [1] designed for testing the saliency versus template hypothesis. Legally and illegally stressed pseudo-words were presented in a passive oddball paradigm. We recorded ERPs in two conditions: the standard and deviant stimuli were simply reversed in order to see if only the absolute differences lead the discrimination or the relative pattern was also taken into account.

2.1. Participants

A total of 60 infants were recruited for the experiment, 48 were included in the statistical analyses. 12 infants were excluded due to extensive artifacts. The recordings were taken at the mean age of 196 days (*SD*=13) in the 6 month-olds' group and 316 days old (*SD*=13) in the 10 month-olds' group. Mean GA was 39.2 weeks while mean birth weight was 3346 g with 548 g standard deviation, with no significant difference between the two age groups. All infants were born to monolingual families and raised in a monolingual environment. None of them had known hearing problems, neurological impairments or any known developmental delay. Parents gave written consent for their child's participation after having detailed information. The experiment was approved by the Ethical Review Committee for Research in Psychology.

2.2. Stimuli

Two types of stimuli, stress variants of a Hungarian pseudo-word ('bebe') with duration of 539 ms, were used. The two stimulus types differed only in their stress pattern. In Hungarian stress is always on the first syllable [21], so for the legal version stress was on the first syllable ('BEbe'). The illegal version was created by reversing the order of the two syllables ('beBE') in order to minimize the difference between the two pseudo-words apart from stress pattern (using Praat [22]). The two syllables were spoken digitalized utterances that differed in three features: maximum intensity (2.42 dB), maximum f0 (15.77 Hz) and rise time (16 ms) respectively, and in contrast to former studies, not in duration. The source stimulus was uttered by a native female speaker in sentence context and edited for computing its illegal variation (for further details read Honbolygó and Csépe [10]).

2.3. Procedure

The procedure was the same as in the Honbolygó and Csépe [1] study. Stimuli were presented in a passive oddball paradigm in random order (deviant probability of 20%). In order to avoid rhythmic affects stimulus onset asynchrony (SOA) varied randomly between 730 and 830ms. We used two stimulus presentation conditions:

• *Legal standard* condition: the legal stimulus was the standard and the illegally stressed pseudo-word was the deviant stimulus.

• *Illegal standard* condition: the stimulus positions were changed, the illegal stimulus was presented as the standard and the legal one was the deviant stimulus.

The order of the two conditions was counterbalanced across subjects. Each condition contained 100 deviants in two blocks. The adult version of the experiment was shortened in order to adapt it to the infant participants [23]. Recording lasted approximately 12 minutes, in order to avoid fatigue among the infants. Stimuli were presented via loudspeaker (Soundkey MS-310, 70 *dB*) that was placed at the distance of 100 cm from the subjects. The experiment was performed using Presentation software (version 12.1, http://www.neurobs.com). The total experimental time was 1 hour including preparation and pauses. Infants were sitting on their parents' lap and were kept calm by presenting cartoons and puppets silently by an assistant.

2.4. Data collection

The EEG was recorded at 500 Hz sampling rate with Ag-AgCl electrodes using an appropriate sized electrode cap (BrainVision Recorder, BrainAmp amplifier, BrainProducts GmbH). Electrodes were attached to F3, Fz, F4, C3, C4, T3, T4, P3, Pz, P4, O1, O2, M1, M2, the reference electrode was Cz. Ground was placed between Fz and Fpz on the midline. The electrode locations corresponded to the international 10-20 system. Offline data analyses were done with BrainVision Analyzer software (BrainProducts GmbH). Recordings were re-referenced to the average reference of M1 and M2 and were band-pass filtered (0,5-20 Hz, 24 dB/oct). Raw EEG data were segmented into 800 ms epochs, including a 100 ms pre-stimulus baseline. Electrophysiological responses to deviant stimuli and to standard appearing right before the deviant were taken into analysis. EEG responses exceeding ±150μV within a sliding window of 300 ms in any channel were rejected automatically. 12 infants' data were rejected because of contaminated recordings. Statistical analyses were carried out in 21 six-month-old and 27 ten-month-old infants for two time windows (300-400 ms, 450-550 ms) based on the grand averages. Epochs were averaged separately for each condition, electrode and participant for all deviants, and for the preceding standards.

3. Results

Difference waves computed by subtracting the averaged standard responses from the deviants are displayed on Figure 1.

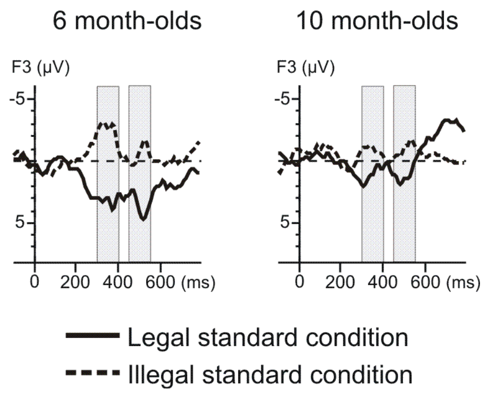

Figure 1: *Difference waves shown for the F3 electrode in two experimental conditions. Responses obtained in the legal standard condition are shown with thick lines, in the illegal standard condition with dotted lines. The time windows -– marked by grey bars show the latency ranges (300-400ms and 450-550ms) where the statistical analysis was run.*

For statistical analysis average areas were computed in the two time windows (300-400 ms; 450-500 ms – right after the syllables). We analyzed the main effect of the conditions performing a 3×2×2 mixed ANOVA with within factors Electrodes (F3, Fz, F4) and Stimulus (standard vs. deviant) and Age (6 MO vs. 10 MO) as a grouping variable in both legal standard and illegal standard conditions. Because of a possible violation of the sphericity assumption we used Greenhouse-Geisser (G-G) adjusted univariate tests where it was necessary [24]. Main effects are reported, effect size were calculated.

In the legal standard condition a Stimulus main effect was revealed confirming the presence of a positive mismatch response (MMR) in both time windows (300-400 ms: $F(1,46)=4.96$, $p<.05$, $r=.31$; 450-550 ms: $F(1,46)=4.76$, $p<.05$, $r=.31$). No Age effect ($p=.76$), or any other significant effect was found.

The statistical analysis did not reveal any significant effect in the illegal standard condition.

4. Discussion

Different ERPs were obtained in the different conditions (role as standard or deviant changed) with the two stimuli. While in the legal standard condition two MMRs were seen, though without any age difference, no significant difference was found in the illegal standard condition. If the responses were based on simple salient acoustic processing only, the automatic change detection reflected by the MMRs would rely on short-term (often called sensory) trace of the stimuli contributing to very similar responses in the two conditions, as the physical differences to detect were the same In contrast, repeating the legal stress pattern as standard enhanced the detection accuracy, so that the illegal pattern used as deviant could activate both a comparison with the long-term representation and with the actual trace of the salient acoustic feature of the stressed initial syllable Moreover, participants had difficulties in detecting stimuli when the illegal version served as standard and the legal one was contrasted with it. In the illegal standard condition we repeated a pattern that contrasted with the long-term representation even if only a weak template has emerged yet, so contrasting it with the legal form as deviant could rely neither on saliency nor on template. Therefore, we interpret our MMR data as electrophysiological evidence of a delicate developmental stage where saliency and emerging stress template representation show a particular dynamics in the first year of life.

Conclusion

Fitting in with the former electrophysiological results in case of other stress- and syllable-timed languages our results have strengthen the view that language specific environmental cues strongly influence the early sensitivity to suprasegmental patterns. Hungarian is a syllable-based language as French, with a word-initial stress pattern as German. Hungarian infants hear trochaic pattern more often, their result are in line with the German infants. Our results imply the existence of a possibly universal developmental pattern of rule abstraction. Relying on rhythmic segmentation procedure infants extract language-specific patterns from their environment, and start to generate expectations. The exact timing of the emergence of a long-term representation is still not clear, although the development of this processing seems to occur in a rather early stage of language acquisition. However, the crucial features of the assumed template as well as their interference with other acoustic attributes of spoken utterances are still unclear. There is a great difference of processing the cues depending on the native language characteristics. According to earlier results, stress based languages provide good conditions for early discrimination of stress, however there are controversial differences regarding syllable based languages [25]. Further investigational question is the nature of long-term stress representation in case of syllable-based languages in accordance with earlier results [26, 27].

5. Acknowledgements

We thank to our colleagues, Gabi Baliga, Kinga Kreif, Orsolya Kolozsvári, Ágoston Török and Andrea Kóbor for their valuable contribution as well to the participating families.

This study was supported by the Hungarian Research Fund project (OTKA-NK No. 101 087, PI: Valéria Csépe).

6. References

[1] Gleitman, L. and Wanner, E. (1982). “Language acquisition: The state of the state of the art”. In E. Wanner and L. Gleitman [Ed], Language Acquisition: The state of the art, 3-48, Cambridge University Press, 1982.

[2] Christophe, A., Dupoux, E., Bertoncini, J. and Mehler, J., “Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition”, J. Acoust. Soc. Am., 95(3):1570–1580, 1994.

[3] Gervain, J. and Mehler, J., “Speech perception and language acquisition in the first year of life”, Annu. Rev. Psychol., 61:191–218, 2010.

[4] Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J. and Werker, J. F., “The development of perceptual grouping biases in infancy: a Japanese-English cross-linguistic study” Cognition, 115:356–361, 2010.

[5] Gervain, J., Nespor, M., Mazuka, R., Horie, R. and Mehler, J., „Bootstrapping word order in prelexical infants: a Japanese-Italian cross- linguistic study”, Cogn. Psychol., 57:56-74, 2008.

[6] Christiansen, M. H., Allen, J. and Seidenberg, M. S., „Learning to segment speech using multiple cues: A connectionist model” Lang. Cognitive Proc., 13:221–268, 1998.

[7] Nazzi, T., Bertoncini, J. and Mehler, J., “Language discrimination by newborns: Toward an understanding of the role of rhythm”, J. Exp.l Psychol: Human Perception and Performance”, 24(3): 756–766, 1998.

[8] Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W. and Jusczyk, A. M., “Six month olds’ detection of clauses embedded in continuous speech: effects of prosodic well-formedness”, Infancy 1: 123–147., 2000.

[9] Nazzi, T., Iakimova, G, Fredonie, S. and Alcantara, C., “Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences”, J. Mem. Lang, 54(3): 283–299., 2006.

[10] Honbolygó, F. and Csépe, V., “Saliency or template? ERP evidence for long-term representation of word stress”, Int. J. Psychophysiol., 87:165-172, 2013.

[11] Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A. and Näätänen, R., “Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations”, Cogn. Brain Res., 73:357-69, 1999.

[12] Winkler, I., Denham, S. L. and Nelken, I., „Modeling the auditory scene: predictive regularity representations and perceptual objects”, Trends Cogn. Sci., 1312:532-540, 2009.

[13] Friederici, A. F., Friedrich, M. and Christophe, A., “Brain responses in 4-month old infants are already language specific”, Curr. Biol., 17:1208–1211, 2007.

[14] Kushnerenko, E., Ceponiene, R., Fellman, V., Huotilainen, M. and Winkler, I., „Event-related potential correlates of sound duration: similar pattern from birth to adulthood”, NeuroReport, 12(17):3777-3781, 2001.

[15] Trainor, L. J. and Adams, B., “Infants’ and Adults' use of duration and intensity cues in the segmentation of tone patterns”, Percept Psychophys., 62: 333–340., 2000.

[16] Leppänen, P. H. T., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M. and Lyytinen, H., “Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia”, Dev. Neuropsychol., 22:407-422, 2002.

[17] Friedrich, M., Herold, B. and Friederici, A.D., “ERP correlates of processing native and non-native language word stress in infants with different language outcomes”, Cortex, 45:662–676, 2009.

[18] Weber, C., Hahne, A., Friedrich, M., Friederici and A. D., „Discrimination of word stress in early infant perception: electrophysiological evidence”, Cogn. Brain Res., 18:149-161, 2004

[19] Cutler, A. and Foss, D. J.,“On the role of sentence stress in sentence processing”, Lang. Speech, 20: 1–10, 1977.

[20] Cutler, A. and Norris, D., “The role of strong syllables in segmentation for lexical access”, Journal of J. Exp. Psychol.: Hum. Percept. Perform, 14(1):113–121, 1988.

[21] Siptár, P., Törkenczy, M., The phonology of Hungarian. Oxford University Press, 2007.

[22] Boersma, P. and Weenink, D., “Praat: doing phonetics by computer”, retrieved from http://www.praat.org/, 2007.

[23] Rivera-Gaxiola, M. and Silva-Pereyra, J., Kuhl, P. K., “Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants”, Dev. Sci., 8:2, 162–172, 2005.

[24] Greenhouse, S. W. and Geisser, S., “On methods in the analysis of profile data”, Psychometrika, 242:95-112, 1959.

[25] Thiessen, E. D. and Saffran, J. R., “When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants.” Dev. Psychol., *39*(4): 706–716, 2003.

[26] Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J. and Nazzi, T., “Language specific prosodic preferences during the first half year of life: evidence from German and French infants.” Infant Behav. Dev., 32:262–74, 2009.

[27] Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Limissuri L. A. and Peperkamp, S., “Language-specific stress perception by 9-month-old French and Spanish infants.” Developmental Sci., 12(6):914–919, 2009.