Language-specific speech perception as mismatch negativity in 10-month-olds' ERP data

Iris-Corinna Schwarz[1], Malin Forsén[2], Linnea Johansson[2], Catarina Lång[2], Anna Narel[2], Tanya Valdés[2], and Francisco Lacerda[1]
[1]Department of Linguistics, Stockholm University
[2]Department of Clinical Science, Intervention, and Technology, Speech Pathology Division, Karolinska Institutet, Stockholm

Abstract

Discrimination of native and nonnative speech contrasts, the heart of the concept of language-specific speech perception, is sensitive to developmental change in speech perception during infancy. Using the mismatch negativity paradigm, seven Swedish language environment 10-month-olds were tested on their perception of six different consonantal and tonal Thai speech contrasts, native and nonnative to the infants. Infant brain activation in response to the speech contrasts was measured with event-related potentials (ERPs). They show mismatch negativity at 300 ms, significant for contrast change in the native condition, but not for contrast change in the nonnative condition. Differences in native and nonnative speech discrimination are clearly reflected in the ERPs and confirm earlier findings obtained by behavioural techniques. ERP measurement thus suitably complements infant speech discrimination research.

Introduction

Speech perception bootstraps language acquisition and forms the basis for later language development. During the first six months of life, infants are 'citizens of the world' (Kuhl, 2004) and perform well in both nonnative and native speech discrimination tasks (Burnham, Tyler, & Horlyck, 2002).

For example, 6-month-old English and German infants tested on a German, but not English contrast [dut]-[dyt], and an English, but not German contrast [dɛt]-[dæt], discriminated both contrasts equally well (Polka & Bohn, 1996). Around the age of six months, a perceptual shift occurs in favour of the native language, earlier for vowels than for consonants (Polka & Werker, 1994). Around that time infants' nonnative speech discrimination performance starts to decline (Werker & Lalonde, 1988), while they continue to build their native language skills (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). For example, 10- to 12-month-old Canadian English environment infants could neither discriminate the nonnative Hindi contrast [ʈɑ]-[tɑ] nor the nonnative Thompson[1] contrast [k̀i]-[q̀i] whereas their 6- to 8-month-old counterparts still could (Werker & Tees, 1984).

This specialisation in the native language holds around six months of age even on a supra-segmental language level: American English language environment infants younger than six months are equally sensitive to all stress patterns of words, and do not only prefer the ones predominantly present in their native language as infants older than six months do (Jusczyk, Cutler, & Redanz, 1993).

During the first year of life, infants' speech perception changes from language-general to language-specific in several features. Adults are already so specialised in their native language that their ability to discriminate nonnative speech contrasts is greatly diminished and can only partially be retrained (Tees & Werker, 1984; Werker & Tees, 1999, 2002). By contrasting native and nonnative discrimination performance the degree of language-specificity in speech perception is shown and developmental change can be described (Burnham, 2003). In the presence of experience with the native language, language-specific speech perception refines, whereas in the absence of experience nonnative speech perception declines. This study focuses on 10-month-olds whose speech perception is language-specific.

A common behavioural paradigm used to test discrimination abilities in infants younger than one year is the conditioned head-turn method (e.g., Polka, Colantonio, & Sundara, 2001). This method requires the infant to be able to sit on the parent's lap and to control head movement. Prior to the experimental test phase, a training phase needs to be incorporated into the experiment to build up the association between perceived changes in contrast presentation and reward display in the infants. The number of trials that it takes the infant to reach criterion during training significantly reduces the possible number of later test trials since the total test time of 10 min maximum remains invariant in infants.

Can electroencephalography (EEG) measurement provide a physiological correlate to the behavioural discrimination results? The answer is yes. Brain activation waves in response to stimulus presentation are called event-related potentials (ERPs) and often show a stimulus-typical curve (Teplan, 2002). This can be for example a negativity response in the ERP, called mismatch negativity (MMN), reflecting stimulus change in a series of auditory signals (Näätänen, 2000). MMN reflects automatic change detection processes on neural level (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002). It is also used in neonate testing as it is the earliest cognitive ERP component measurable (Näätänen, 2000). The general advantage of using ERPs in infant research lies exactly within in the automaticity of these processes that does neither demand attention nor training (Cheour, Leppänen, & Kraus, 2000).

For example, mismatch negativity represents 6-month-olds' discrimination of consonant duration changes in Finnish non-sense words (Leppänen et al., 2002). Similarly, differences in the stress patterns of familiar words are reflected in the ERPs of German and French 4-month-olds (Friederici, Friedrich, & Christophe, 2007). This reveals early language-specific speech perception at least in suprasegmental aspects of language.

How infant language development from language-general to language-specific discrimination of speech contrasts can be mapped onto neural response patterns was demonstrated in a study with 7- and 11-month-old American English environment infants (Rivera-Gaxiola, Silva-Pereya, & Kuhl, 2005). The infants could be classed into different ERP patterns groups, showing not only negativity at discrimination but also positive differences. Discrimination of Spanish voice-onset time (VOT) differences was present in the 7-month-olds but not in the 11-month-olds (Rivera-Gaxiola et al., 2005).

Hypothesis

If ERPs and especially mismatch negativity are confirmed by the current study as physiological correlates to behavioural infant speech discrimination data, 10-month-old Swedish language environment children would discriminate native, but not nonnative contrast changes, as they should perceive speech in a language-specific manner at this stage of their development.

Method

Participants

Seven 10-month-old infants (four girls and three boys) participated in the study. Their average age was ten months and one week, with an age range of ten to eleven months. The participants' contact details were obtained from the governmental residence address registry. Families with 10-month-old children who live in Greater Stockholm were randomly chosen and invited to participate via mail. They expressed their interest in the study by returning a form on the basis of which the appointment was booked over the phone. All children were growing up in a monolingual Swedish-speaking environment. As reward for participation, all families received certificates with a photo of the infant wearing the EEG net.

Stimuli

Speech stimuli were in combination with the vowel /a/ the Thai bilabial stops /b̬/, /b/, and /p^h/ and the dental/alveolar plosives /d̬/, /d/, and /t^h/ in mid-level tone (0), as well as the velar plosive [ka] in low (1), high falling (2), and low rising (4) tone. Thai distinguishes three voicing levels. In the example of the bilabial stops this means that /b̬/ has a VOT of -97 ms, /b/ of 6 ms and /p^h/ of 64 ms (Burnham, Francis, & Webster, 1996). Out of these three stimulus sets contrast pairs were selected that can be contrastive (native) or not contrastive (nonnative) in Swedish. The consonantal contrasts [ba]-[p^ha] and [da]-[t^ha] are contrastive in Thai and in Swedish, whereas the consonantal contrasts [b̬a]-[ba] and [d̬a]-[da] are only contrastive in Thai.

Both consonantal contrasts were mid-tone exemplars but the third set of contrasts was tonal. It presents the change between high falling and low rising tone in the contrast [ka2]-[ka4]

and between low and low rising tone in the contrast [ka1]-[ka4]. Although the two tonal contrasts must be considered nonnative to Swedish infants, non-Thai listeners in general seem to rely on complex acoustic variables when trying to discriminate tone (Burnham et al., 1996), which therefore makes it difficult to predict the discrimination of tonal contrast change.

After recording, all speech stimuli were presented to an expert panel consisting of two Thai native speakers and one trained phonetician in order to select the three best exemplars per stimulus type out of ten (see Table 1 for variation of utterance duration between the selected exemplars).

Table 1. The table shows utterance duration in ms for all selected exemplars per stimulus type and average duration in ms per stimulus type (Thai tone is demarcated by number).

	b̬a0	ba0	pʰa0	da0	d̬a0	taʰ0	ka1	ka2	ka4
1	606	576	667	632	533	607	613	550	535
2	626	525	646	634	484	528	558	502	538
3	629	534	629	599	508	585	593	425	502
M	620	545	647	622	508	573	588	492	525

Within each trial, the first contrast of each pair was repeated two to five times until the second contrast was presented twice after a brief interstimulus interval of 300 ms. Each stimulus type (native consonantal, nonnative consonantal, nonnative tonal) was presented twelve times within a block. Within each block, there were 36 change trials and nine no-change trials. A change trial repeated identical exemplars for the first contrast and then presented the identical exemplar of the second contrast twice. A no-change trial had identical first and second sound exemplars, presented randomly between four and seven times. A completed experiment consisted of three blocks à 145 trials.

Equipment

The EEG recordings took place in a radiation-insulated near-soundproof test chamber at the Phonetics Lab at Stockholm University.

Infant brain activation was measured by EGI Geodesic Hydrocel GSN Sensor nets with 124 electrodes on the infant net sizes. These net types permit EEG measurement without requiring gel application which makes them particularly compatible with infant research; potassium chloride and generic baby shampoo serve as conductive lubricants instead. All electrode impedances were kept below 50 kΩ at measurement onset. All EEG channel data was amplified with an EGI NetAmps 300 amplifier and recorded with a sampling rate of one sample every 4 ms. The program Netstation 4.2.1 was used to record and analyse the ERPs.

The stimuli were presented with KOSS loudspeakers, mounted at a distance of about 100 cm in front of the child. The volume was set to 55 dB at the source. The experiment was programmed and controlled by the e-prime 1.2 software.

Procedure

All infant participants were seated in their parent's lap, facing a TV screen on which silenced short cartoon movie clips played during the experiment to entertain the infants and keep them as motionless as possible. The infants were permitted to eat, breastfeed, sleep, as well as suck on dummies or other objects during stimulus exposure.

Dependent on the randomisation of the first contrast between two and five repetitions, the duration of the entire experiment varied between 10 and 13 min. Infant and parent behaviour was monitored through an observer window and the experiment was aborted in the case of increasing infant fussiness - this happened in one case after 8 min of stimulus exposure.

Data treatment

The EEG recordings were filtered with a band-pass filter of 0.3 to 50 Hz and clipped into 1000 ms windows starting at the onset of the second contrast. These windows were then cleaned from all 10 ms segments during which the ERP curve changed faster than 200 µV to remove measurement artefacts caused by body movement and eye blinks. If more than 80% of the segments of one single electrode were marked as artefacts, the entire data from that electrode was not included in the average.

Results

In accordance with other infant speech perception ERP studies (e.g., Friederici et al., 2007), the international 10-20 electrode system was selected to structure the EEG data. Within this system, the analysis focused on electrode T3, situated at the temporal lobe in the left hemisphere, as MMN in 8-month-old infants has previously been found to be largest in T3 (Pang et al., 1998).

Comparing native and nonnative consonantal contrast change trials, the curve for [ba]-[p^ha] and [da]-[t^ha] that are contrastive both in Thai and in Swedish shows a dip between 200 and 350 ms (Figure 1).

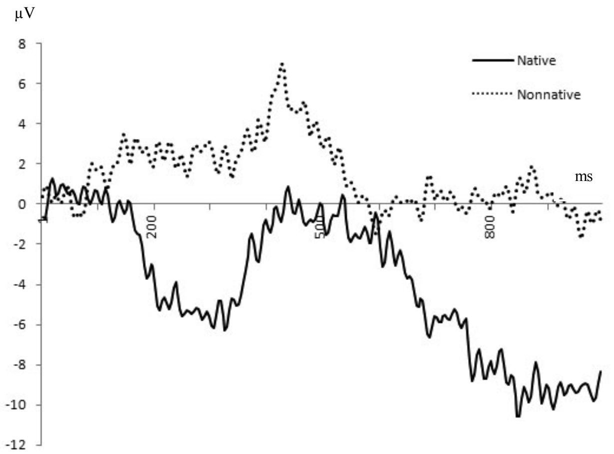

Figure 1. The graph compares the ERPs for the native and nonnative consonantal contrast change trials during 1000 ms after stimulus onset of the second contrast. The ERP for the native condition shows mismatch negativity in µV between 200 and 350 ms, typical for the discrimination of auditory change. The ERP for the nonnative condition however only shows a stable continuation of the curve.

This negativity response is mismatch negativity and at the same time a sign that the 10-month-olds discriminated the contrast change. It peaks at 332 ms with -6.3 µV.

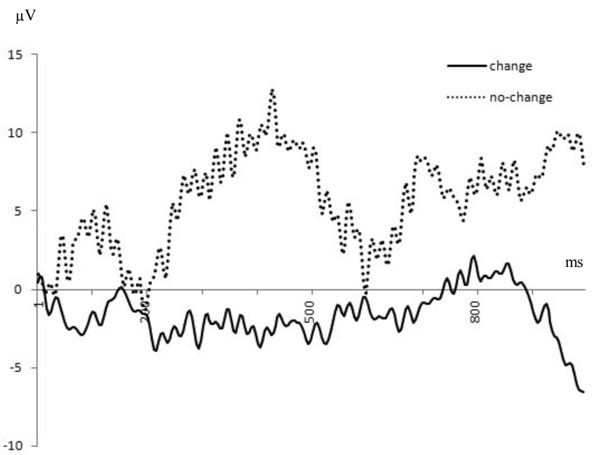

Figure 2. The graph shows the ERPs for the nonnative tonal change and no-change trials in µV during 1000 ms after stimulus onset of the second contrast (which is of course identical to the first in the no-change condition). No mismatch negativity can be observed in either condition.

However, the curve for [b̬a]-[ba] and [d̬a]-[da] that are contrastive in Thai, but not in Swedish, shows no negativity response. The infants were not able to detect the contrast change in the nonnative consonantal condition, as the flat graph indicates. The two curves differ significantly between 200 and 350 ms ($p<.001$), demonstrating that the neural responses of 10-month-old Swedish language environment infants discriminate native but not nonnative consonantal contrast changes. These discrimination abilities show in the typical neural mismatch negativity response.

Comparing the nonnative tonal contrast change condition to its no-change condition, the previous result of a lack of discrimination for nonnative contrasts in 10-month-olds is regenerated (Figure 2). The graph for the nonnative tonal contrast change remains relatively flat and stable throughout the early window typical for mismatch negativity responses, while being negative however. The no-change curve for the tonal condition on the other hand responds with increasing positive neural activation to the repetition of the first contrast that peaks just after 400 ms with 12.8 µV. Neither of the tonal conditions shows mismatch negativity.

Discussion

This study shows that ERPs and in particular the concept of mismatch negativity reflect reliably infant speech discrimination abilities, previously mostly demonstrated in behavioural experiments, and confirms ERP data as a physiological correlate to those.

It replicated the findings of Rivera-Gaxiola and colleagues in 11-month-old American infants discriminating English, but not Spanish contrasts (2005) with Swedish 10-month-olds. The Swedish infant participants discriminated only Thai contrasts that are also legitimate in Swedish, but not those that are illegitimate in Swedish. This result is in line with our prediction that contrasts that sound native would be discriminated, but not those considered to be nonnative. At ten months, infants' speech perception is language-specific, which results in good discrimination abilities of native speech sounds and a loss of discrimination abilities of nonnative speech sounds. This establishes a neural basis to the well-known findings from conditioned head-turn studies for example.

On a side note: in order to be able to truly speak of a decline or a loss of the nonnative speech discrimination abilities in the Swedish

10-month-old infants, further studies with infants younger than six months are necessary and currently under way to provide the required developmental comparison.

MMN was strongest in the nonnative consonantal change condition in this study. Even though the tonal stimuli are generally very interesting for infants and potentially not processed the same way as speech, the absence of clear mismatch negativity shows that the 10-month-olds' brains did not react in the same way to change in tones as they did to change in consonants in the native condition. Furthermore, repetition of the same tone elicited higher absolute activation than change in a series of tonal speech sounds.

Interestingly, MMN is in this study the only response pattern to a detected change in a series of speech sounds. Rivera-Gaxiola and colleagues had found subgroups with positive or negative ERP discrimination curves in 11-month-olds (2005), whereas MMN is a strong and stable indicator of discrimination in our participants. And this is the case even with varying repetition distribution of the first contrast taking up between 50 % and 70 % of a trial (corresponding to two to five times) in comparison to the fixed presentation of the second contrast (two times). Leppänen and colleagues (2002) presented the standard stimulus with 80% and each of their two deviant stimulus with 10 % probability with a 610 ms interstimulus interval observe MMN in 6-month-olds. The MMN effect is therefore quite robust to changes in trial setup, at least in 10-month-olds.

Conclusions

The ERP component mismatch negativity (MMN) is a reliable sign for the detection of change in a series of speech sounds in 10-month-old Swedish language environment infants. For the consonantal contrasts, the infants' neural response shows discrimination for native, but not nonnative contrasts. Neither do the infants indicate discrimination of the nonnative tonal contrasts. This confirms previous findings (Rivera-Gaxiola et al., 2005) and provides physiological evidence for language-specific speech perception in 10 month-olds.

Acknowledgements

This study is the joint effort of the language development research group at the Phonetics Lab at Stockholm University and was funded by the Swedish Research Council (VR 421-2007-6400), the Knut and Alice Wallenberg Foundation (Grant no. KAW 2005.0115), and the Bank of Sweden Tercentenary Foundation (MILLE, RJ K2003-0867), the contribution of all of which we acknowledge gratefully. We would also like to thank all participating families: without your interest and dedication, our research would not be possible.

Footnotes

[1] *Thompson is an Interior Salish (Native Indian) language spoken in south central British Columbia. In native terms, it is called Nthlakampx or Inslekepmx. The example contrast differs in place of articulation.*

References

Burnham, D. (2003). Language specific speech perception and the onset of reading. *Reading and Writing: An Interdisciplinary Journal, 16*(6), 573-609.

Burnham, D., Francis, E., & Webster, D. (1996). *The development of tone perception: Cross-linguistic aspects and the effect of linguistic context.* Paper presented at the Pan-Asiatic Linguistics: Fourth International Symposium on Language and Linguistics, Vol. 1: Language and Related Sciences, Institute of Language and Culture for Rural Development, Mahidol University, Salaya, Thailand.

Burnham, D., Tyler, M., & Horlyck, S. (2002). Periods of speech perception development and their vestiges in adulthood. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 281-300). Trier: Wissenschaftlicher Verlag.

Cheour, M., Leppänen, P. H. T., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111*(1), 4-16.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language-specific. *Current Biology, 17*(14), 1208-1211.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*(3), 675-687.

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

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606-608.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: A longitudinal ERP study. *NeuroReport, 13*(15), 1843-1848.

Leppänen, P. H. T., Richardson, U., Pihko, E., Eklund, K., Guttorm, T. K., Aro, M., et al. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Developmental Neuropsychology, 22*(1), 407-422.

Näätänen, R. (2000). Mismatch negativity (MMN): perspectives for application. *International Journal of Psychophysiology, 37*(1), 3-10.

Pang, E. W., Edmonds, G. E., Desjardins, R., Khan, S. C., Trainor, L. J., & Taylor, M. J. (1998). Mismatch negativity to speech stimuli in 8-month-olds and adults. *International Journal of Psychophysiology, 29*(2), 227-236.

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America, 100*(1), 577-592.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109*(5 Pt 1), 2190-2201.

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

Rivera-Gaxiola, M. C. A., Silva-Pereya, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*(2), 162-172.

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

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review, 2*(2), 1-11.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*(5), 672-683.

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

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535.

Werker, J. F., & 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.