Neuroscience Letters 651 (2017) 237–241

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

Mismatch negativity (MMN) to speech sounds is modulated systematically by manual grip execution

Mikko Tiainen [a,*], Kaisa Tiippana [a], Petri Paavilainen [a,b], Martti Vainio [c], Lari Vainio [a]

[a] *Department of Psychology and Logopedics, University of Helsinki, Finland*
[b] *Cognitive Brain Research Unit, University of Helsinki, Finland*
[c] *Department of Modern Languages, University of Helsinki, Finland*

HIGHLIGHTS

- The influence of grip actions on the MMNs to heard syllables was examined.
- MMN was smaller to a syllable congruent with the grip than to an incongruent one.
- This shows for the first time that manual actions systematically modulated the MMN.
- Grasping and speech interact at the pre-attentive level, reflected by the MMN.

ARTICLE INFO

Article history:
Received 8 March 2017
Received in revised form 21 April 2017
Accepted 10 May 2017
Available online 11 May 2017

Keywords:
MMN
Speech
Action
Action-perception
Gestures

ABSTRACT

Manual actions and speech are connected: for example, grip execution can influence simultaneous vocalizations and vice versa. Our previous studies show that the consonant [k] is associated with the power grip and the consonant [t] with the precision grip. Here we studied whether the interaction between speech sounds and grips could operate already at a pre-attentive stage of auditory processing, reflected by the mismatch-negativity (MMN) component of the event-related potential (ERP). Participants executed power and precision grips according to visual cues while listening to syllable sequences consisting of [ke] and [te] utterances. The grips modulated the MMN amplitudes to these syllables in a systematic manner so that when the deviant was [ke], the MMN response was larger with a precision grip than with a power grip. There was a converse trend when the deviant was [te]. These results suggest that manual gestures and speech can interact already at a pre-attentive processing level of auditory perception, and show, for the first time that manual actions can systematically modulate the MMN.

1. Introduction

Hand and mouth movements are interconnected. For example, while pronouncing a syllable, the mouth is opened wider when watching a large object being grasped than when watching a small object being grasped [1]. In addition, the so-called mouth–hand mimicry theories suggest that people tend to mimic concurrent hand movements with their mouth [2,3]. Indeed, young children [4], and also chimpanzees [5], have a tendency to perform mouth movements in an imitative manner with fine-motor hand movements.

We have previously shown that speech sounds and hand grips have specific connections (e.g. [6,7]). For example, the consonant [k] is associated with the power grip and the consonant [t] with the precision grip, so that grip execution is faster when an associated consonant is simultaneously pronounced or prepared. This congruency effect can also be observed in vocal reaction times [7]. Our findings provide evidence of the systematic nature of the connections between hand and mouth movements, and are thus in line with the mouth-hand mimicry theories as well as with the theories proposing that manual gestures might have contributed to the development of articulatory gestures (e.g. [8,9]). We have hypothesized that the articulations involving the tongue body, such as [k] could be considered an analogue of a power grip, where an object is grasped by pressing it against the palmar surface of the hand (e.g. [10]). In contrast, [t], for which the tip of the tongue is used, would be the analogue of a precision grip, where the tips of the index finger and thumb are utilized to pick up small objects [10].

* Corresponding author at: Siltavuorenpenger 1A, Room 427, 00014 University of Helsinki, Finland.
E-mail address: mikko.o.tiainen@helsinki.fi (M. Tiainen).

http://dx.doi.org/10.1016/j.neulet.2017.05.024
0304-3940/

We have also shown that the above-discussed articulation-grip connection works in a syllable categorization task [11]. In this study, participants were asked to prepare a power or a precision grip and then observe a noise-masked syllable. They executed the grip at the end of the syllable presentation and then reported whether the presented syllable was [ke] or [te]. When participants prepared/executed a power grip, they were more likely to categorize the presented syllable as [ke], and when a precision grip was prepared/executed, syllables were more often categorized as [te]. We found evidence that this categorization effect operates, at least in part, at the decision-making stage, but possible influences at earlier stages of perceptual processing could not be addressed thoroughly. In fact, it has been shown that stimulating mouth motor areas with transcranial magnetic stimulation (TMS) modulates speech perception systematically [12–14]. Therefore, if mouth and hand actions share overlapping processing networks, it is likely that the modulation of speech categorization by grips could also operate at early perceptual processing stages, in addition to the decision-making stage. This idea is further supported by a recent functional magnetic resonance imaging (fMRI) study, which showed that grip execution modulates the activity in the auditory cortex while listening to meaningless non-speech sounds [15].

In the current event-related potential (ERP) study, we explored whether precision and power grip performance could have a systematic influence on perceptual processes related to heard syllables by utilizing the well-known ERP component, the mismatch negativity (MMN, e.g. [16]). The MMN is usually elicited in an oddball paradigm where occasional deviant stimuli (e.g. a tone of a different frequency) are presented in the midst of a stimulus sequence consisting of repetitive, physically identical “standard” stimuli. The MMN is seen in the deviant minus standard difference curve as a negativity peaking at around 100–200 ms after the deviant-stimulus onset. The MMN originates from the auditory cortex and is related to early stimulus processing (e.g. [17]). It has been theorized that representations of the regularities of auditory stimuli are stored in the auditory cortex and compared to incoming auditory information. The MMN represents the mismatch between the predicted regularity and the actual input [18]. As MMN is observed even though no attention is paid to the auditory information, it is considered to reflect early and relatively automatic stages of auditory processing [16]. The MMN is also observed when using speech stimuli [16], and even with the McGurk illusion [19–21]. In the McGurk illusion, presenting incongruent visual and auditory speech stimuli alters the auditory perception [22]. These studies show that visual information can influence auditory processing at the pre-attentive level [19–21].

As the MMN to speech stimuli can be modulated by stimuli of another modality (i.e. visual speech), the question arises whether it could be influenced by other non-auditory means as well. As our previous study showed that grip actions can influence speech categorization [11], we hypothesized that grip execution could also specifically modulate MMN to speech stimuli when the grips are associated with the speech stimuli. If this modulation could be observed, it would show for the first time that action-related processes modulate MMN patterns, and it would also strongly support the notion that manual actions interact with speech processing at an early stage and thus can influence the perceptual outcome. We used syllables [te] and [ke] as stimuli since they have been shown to be robustly associated with the precision and power grip, respectively [6,11]. We employed a classic oddball paradigm, where one syllable (e.g. [te]) was the standard stimulus, and the other one was the deviant (e.g. [ke]). In another block, the setting was reversed. The participants' task was to respond with appropriate grips to onscreen cues, while ignoring the sounds. Since grips were executed both during standard and deviant stimuli, we could create difference curves from trials that had exactly the same stimulus, but the grip context was different (i.e. physically the same speech stimulus as a standard vs. deviant, presented during two different types of grips). Thus, the possible effects on the MMNs could not be simply explained either by the two syllables producing different MMNs or by the brain responses being differentially contaminated by motor activity related to grip execution.

2. Material and methods

2.1. Participants

Twenty-one native Finnish-speaking volunteers participated in the study (19 women, 2 men, aged 23.6 ± 5.8 years). All were right-handed, and reported normal vision, hand motor functioning and hearing. This research was approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences, and the participants gave their written informed consent.

2.2. Equipment, stimuli and procedure

The participant sat in a dimly-lit sound-attenuated room 100 cm in front of a 17″ LCD-screen, holding two grip devices marked with green and blue tape in his/her right hand. The precision grip device was held between the index finger and thumb, and the power grip device with the rest of the fingers, squeezed against the palm of the hand (for more details on the grip devices see, e.g., Vainio et al., 2013).

The auditory stimuli were [ke] and [te] syllables, recorded from the speech of a Finnish woman and delivered binaurally through headphones. Both syllables were 160 ms long and edited to have an equal average f_0 of 241 Hz The syllables were presented with an intensity of 50 dBA. The visual stimulus was a green or a blue circle (diameter 1.53° of visual angle) shown in the centre of the screen.

The experiment used an oddball paradigm where one of the syllables was the standard stimulus (P = .85) and the other one was the deviant stimulus (P = .15). The participant's task was to respond to the circles when they appeared on the screen by squeezing the corresponding, colour-matching grip device. The participant was told to ignore the auditory stimuli. The syllables were presented with 500-ms interstimulus intervals. When a grip response was required, the visual go-signal for the response (green or blue circle) appeared on the screen 100 ms before the syllable was presented, so that the actual grip processing could start around the same time as the syllable processing, and the circle remained onscreen for 400 ms. The sequence of the deviant presentation was randomized except that two deviants were never presented in a row, and all grip responses were preceded by a standard stimulus. The experiment was split into two blocks where in one the standard was [ke] and the deviant was [te] and vice versa in the other one. The order of the blocks was balanced across participants, as was the colour matching for the grip responses. Each block had six trial types: standard with no grip (2150 trials), standard with a power grip (300 trials), standard with a precision grip (300 trials), deviant with no grip (150 trials), deviant with a power grip (150 trials) and deviant with a precision grip (150 trials). There were more grip responses with standard stimuli to keep the task of the participant more involving and also to prevent associating grips directly with the deviant stimuli.

In total, one block had 3000 trials. Each block was split into three runs, and each run had three pauses whose length was participant-controlled. Between blocks there was a longer pause in order for the participant to rest properly. One run lasted for about 15 min, depending on the length of the pauses, and the entire experiment with preparations about 2.5 h.