Journal of Motor Behavior

Routledge
Taylor & Francis Group

ISSN: 0022-2895 (Print) 1940-1027 (Online) Journal homepage: https://www.tandfonline.com/loi/vjmb20

# The Influence of Number Magnitude on Vocal Responses

L. Vainio, T. Mustonen & M. Vainio

**To cite this article:** L. Vainio, T. Mustonen & M. Vainio (2019) The Influence of Number Magnitude on Vocal Responses, Journal of Motor Behavior, 51:2, 129-140, DOI: 10.1080/00222895.2018.1440522

**To link to this article:** https://doi.org/10.1080/00222895.2018.1440522

Journal of Motor Behavior, Vol. 51, No. 2, 2019


RESEARCH ARTICLE

OPEN ACCESS

Check for updates

# The Influence of Number Magnitude on Vocal Responses

**L. Vainio[1], T. Mustonen[1], M. Vainio [2]**

[1]Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland. [2]Phonetics and Speech Synthesis Research Group, Department of Modern Languages, University of Helsinki, Finland.

**ABSTRACT.** The study investigated whether number magnitude can influence vocal responses. Participants produced either short or long version of the vowel [ɑ] (Experiment 1), or high or low-pitched version of that vowel (Experiment 2), according to the parity of a visually presented number. In addition to measuring reaction times (RT) of vocal responses, we measured the intensity, the fundamental frequency ($f_0$) and the first and second formants of the vocalization. The RTs showed that the long and high-pitched vocal responses were associated with large numbers, while short and low-pitched vocal responses were associated with small numbers. It was also found that high-pitched vocalizations were mapped with the odd numbers, while the low-pitched vocalizations were mapped with the even numbers. Finally, large numbers increased the $f_0$ values. The study shows systematic interactions between the processes that represent number magnitude and produce vocal responses.

*Keywords*: number magnitude, motor response, vocalization, reaction time

## Introduction

Ability for abstraction is a basic cognitive human skill that allows the semantic representation of relative attributes such as short-long, small-large, and light-heavy, and the utilization of these types of representations in a flexible manner for representing different concepts (e.g., short time/person, small stone/number, light load/feeling). Indeed, people show a functional overlap between the representations of, for example, length and duration (e.g., Lourenco & Longo, 2010; Srinivasan & Carey, 2010). A Theory of Magnitude (ATOM) (Bueti & Walsh, 2009; Walsh, 2003) correspondingly assumes that a common system in the parietal cortex is responsible for processing magnitude information related to space, time and quantity (e.g., numbers, the brightness of a light, or the loudness of a sound).

One central assumption of the ATOM hypothesis is that this common magnitude system ultimately serves action planning; it overlaps with the processes that transform the physical magnitude information of external objects into the corresponding motor responses. For instance, the processes that transform the size of an object into the corresponding motor programs of grasp action—defining which grip type (i.e., precision or whole hand) has to be selected and how wide grip opening is required—are also employed when one has to, for example, estimate the relative magnitudes of viewed numbers. Indeed, it has been shown that grasp performance is similarly facilitated by the magnitude information of the stimulus when participants are required to perform either a precision pinch or whole hand grasp response according to the shape of the graspable object (Ellis, Tucker, Symes, & Vainio, 2007), or the parity of the number (Lindemann, Abolafia, Girardi, & Bekkering, 2007; Moretto & Di Pellegrino, 2008). Small objects and numbers facilitate precision pinch responses, whereas large objects and numbers facilitate whole hand grasp responses. Correspondingly, Andres, Davare, Pesenti, Olivier, and Seron (2004) found that participants produce a grip opening response rapidly when they have to judge the parity of a number whose size is relatively large (e.g., 8 or 9). In contrast, small numbers (e.g., 1 and 2) facilitated grip closure responses. More recently, it has been found that this phenomenon can be also observed in the kinematics of grasp movements (Andres, Ostry, Nicol, & Paus, 2008). When participants are required to reach and grasp a wooden block with a number (1, 2, 8, or 9) printed on the face of it, the grip aperture increases as a function of the number magnitude.

In another research tradition, it has been shown that processes related to controlling different grasp properties are connected to processes that program mouth movements in general and articulatory gestures in parti-cular. Originally, Gentilucci, Benuzzi, Gangitano, and Grimaldi (2001) showed that when participants have to perform a grasp action and simultaneously open their mouth, the more the manual task required grip opening, the more the lip aperture is increased. Correspondingly, the vocalization of an open vowel [ɑ] results in increased grip opening in comparison to vocalizing a closed vowel [i] when the vowel production and grasping are performed simultaneously (Gentilucci & Campione, 2011). In addition, it has been shown that when participants are required to perform either the precision or whole hand grasp response and simultaneously pronounce the vowel as a meaningless speech unit (e.g., [ɑ] or [i]), the grasp and vocal responses are performed relatively rapidly when there is (hypothesized) congruency between the manual and vocal responses (e.g., precision pinch – the close

*Correspondence address: Lari Vainio, Perception, Action & Cognition Research Group, Faculty of Medicine, Department of Psychology and Logopedics, University of Helsinki, Haartmaninkatu 8, PL 63, 00014, Helsinki, Finland. E-mail: lari.vainio@helsinki.fi*

vowel [i]) (Tiainen et al., 2016; Vainio, Schulman, Tiippana, & Vainio, 2013). These findings are in line with the views assuming close neural, functional, developmental and evolutionary connections between manual grasping and articulatory processes (Arbib, 2005; Gentilucci & Corballis, 2006; Rizzolatti & Arbib, 1998).

Importantly for the present purpose, vocal production can be also influenced by visually presented grasp-related information. It has been shown that vowel production is modulated by viewing a graspable object or a hand that is grasping an object. For example, lip opening along with the spectral components of intensity, fundamental frequency ($f_0$) and the first formant (F1) are increased by viewing an object whose size is compatible with the whole hand grasp in comparison to a precision pinch compatible objects (Gentilucci, Campione, Dalla Volta, & Bernardis, 2009). Similarly, Vainio et al., 2017 have shown that participants produce the vowel [i] relatively rapidly when they are presented with an image of a precision pinch whereas the vowel [ɑ] is produced relatively rapidly when they are presented with an image of the whole hand grasp. These findings that associate an increased F1 and the vowel [ɑ] with the whole hand grasp-compatible stimuli are comprehensible if one accepts the link between F1 and vowel openness. In phonetics, vowel openness (height) refers to the aperture of the jaw and the vertical position of the tongue relative to the roof of the mouth. In turn, the height of the tongue is associated with the F1. In close vowels, such as [i], the F1 is consistent with the tongue being positioned relatively high in the mouth (i.e., relatively close to the palate). In contrast, in open vowels, such as [ɑ], F1 is consistent with the jaw being relatively open and the tongue being positioned low in the mouth. It follows from this that the higher the frequency of the first formant is, the more open is the vowel (Ladefoged, 2006). The findings showing association between open vowels, the increased frequency of F1, and whole hand grasp-compatible stimuli suggest that processing visually presented whole hand grasp-compatible stimuli is partially grounded in the articulatory motor mechanisms that automatically increase the opening of the vocal tract when vowel production is required during stimulus processing.

Although number magnitude has been shown to influence grasp responses (Andres et al., 2004), and grasping is tightly linked to processes that program articulation (Gentilucci et al., 2009; Vainio et al., 2013), no previous research, to our knowledge, has explored whether the magnitude of a viewed number can influence the vocal processes. In addition, the ATOM theory makes no clear predictions for the articulatory processes used in speech production and, to date, has focused only on studies using manual responses generated using one's hands or eye gaze (Bueti & Walsh, 2009). In this study, we test the applicability of ATOM to the production of fine motor skills beyond those that are intrinsically spatial, and thereby examine how number magnitude might interact with the processes tied to speech production. Hence, we conducted two experiments to investigate whether the vowel production could be systematically influenced by the magnitude of numbers presented to participants.

## The Research Questions and their Rationale

Most of the studies exploring how the magnitude of viewed numbers influence motor responses have used a parity judgement task in which participants are required to judge whether the number is odd or even by pressing response keys, for example, with their left and right hand. These kinds of studies have shown that right-sided responses are made relatively rapidly when the number is large, whereas left-sided responses are facilitated when the number is small (Dehaene, Bossini, & Giraux, 1993). The effect is observed, for example, when participants are asked to respond using two fingers of the same hand (Kim & Zaidel, 2003) or when a single hand is moved to the right or left side of the initial position of the hand (Fischer, 2003). This effect that was labelled as SNARC (Spatial-Numerical Association of Response Codes) has been taken as an evidence for that the representations of numerical magnitude are coded spatially in a "mental number line."

In this study, we also use a parity judgement task. However, in contrast to the previous investigations, our participants were required to select a specific utterance for the response according to the parity of the number. Similar to some experiments that have explored whether number magnitude influences grasp actions (e.g., Andres et al., 2008), our participants were visually presented with the numbers 1, 2, 8, and 9. The participants were required to judge whether the number is odd or even by producing the vowel [ɑ] in one or another form (e.g., in a short or long form). As already mentioned, the spectral components of intensity, $f_0$ and F1 have been observed to increase when participants are presented with an object whose size is compatible with the whole hand grasp in comparison to a precision pinch-compatible objects (Gentilucci et al., 2009). Similarly, we predicted that these vocal components can be systematically influenced by the size of the number.

Regarding the intensity, the ATOM hypothesis (Walsh, 2003) assumes that overlapping sensory-motor processes are responsible for representing magnitude-related dimensions of loudness and number. Indeed, research has revealed that large numbers and increasing number sequences are automatically associated with relatively loud sounds, whereas small numbers and decreasing number sequences are associated with quieter sounds (Alards-Tomalin, Walker, Nepon, & Leboe-McGowan, 2017; Alards-Tomalin, Walker, Shaw, & Leboe-McGowan, 2015). Also, given that large objects in comparison to small objects have been linked to increased intensity of vocal responses (Gentilucci et al., 2009), we assume that the intensity of a vocalization would be lower when categorizing the numbers 1 and 2 in comparison to the numbers 8 and 9.

Regarding the $f_0$, there are in fact two potential outcomes. On the one hand, it has been shown that small objects are associated with high-pitched sounds while large objects are associated with low-pitched sounds (e.g., Gallace & Spence, 2006; Parise & Spence, 2009). This effect might reflect learned associations between certain auditory and visual features that typically occur together in nature—the larger the object, the lower the frequency (see Coward & Stevens, 2004). On the other hand, Gentilucci et al., 2009 have found that viewing large objects increases the pitch of the vocalization in comparison to small objects. This effect can be assumed to reflect common sensory-motor processes underlying the vowel production and representing magnitude information. This latter observation is in line with previous investigations that associate high-pitched tones with large numbers and low-pitched tones with small numbers. Firstly, Oriet, Tombu, and Jolicoeur (2005) found that participants categorized the magnitude of large numbers (8, 9) faster when the decision was preceded by a high pitch tone, and the magnitude of small numbers (1, 2) faster when the decision was preceded by a low pitch tone. Secondly, in the study reported by Campbell and Scheepers (2015), the participants were required to categorize whether the second number of a pair of sequential auditory numbers was lower or higher in numerical value than the first number. The vocal pitches of the two numbers either ascended or descended. The participants made more errors when the pitch was ascending and the second number was smaller (in magnitude) than the first number, and when the pitch was descending and the second number was larger than the first number. Consequently, it is interesting to observe whether number magnitude can influence $f_0$ values of vocalization, and if it does, whether the results comply with the findings associating small objects with high tones and large objects with low tones (e.g., Coward & Stevens, 2004; Gallace & Spence, 2006; Parise & Spence, 2009), or whether we observe an opposite effect as predicted by the other research tradition (Campbell & Scheepers, 2015; Gentilucci et al., 2009; Oriet et al., 2005).

It is known that F1 component of vocal spectra mostly reflects openness of a vocal tract during vocalization. In general, it has been observed that F1 values are higher for vocalizations that utilize a wider vocal tract (Fant, 1960). Hence, it can be proposed that increase in F1 as a function of an increase in the size of the viewed object, as observed by Gentilucci et al., 2009, is the consequence of automatic modulation of motor planning processes that are triggered by the size of the object. That is, larger objects lead to an increase in the opening of a vocal tract during vocalization, which in turn is observed in higher F1 values. As such, the F1 modulation can be assumed to correspond to the effect of increased finger opening triggered by the viewed numbers of relatively large magnitudes (Andres et al., 2004). In other words, based on the mouth-hand hypothesis discussed above, we predict that viewing large numbers not only increases the grip opening but also the opening of the vocal tract, which in turn results in relatively high F1 values.

In addition to potential modulation of the spectral components, we predicted that the magnitude of a number can influence the latency of vocalization onsets (i.e., vocal reaction times) if there is some congruency between the size of the number and the type of the vocal response. In Experiment 1, the participants were required to produce either short or long version of the vowel [ɑ] according to the parity information of the number, whereas in Experiment 2 they were asked to produce the vowel [ɑ] in low or high pitch according to the parity information. In line with the ATOM hypothesis (Walsh, 2003), a previous study has shown that when participants are asked to press a response key for a short or long duration according to the parity of the number, short responses are produced more rapidly when the number was small, whereas long responses are produced more rapidly when the number was large (Kiesel & Vierck, 2009). Experiment 1 studied whether this effect can be generalized to vocal responses so that short [ɑ]s would be pronounced faster with small rather than large numbers and vice versa for the long [ɑ]s. In addition, we predicted that in Experiment 2 the responses would be produced particularly rapidly when the required response (i.e., low versus high-pitched vocalization) was congruent with the size of the number. Given that previous studies have associated high pitch tones with relatively large numbers and low pitch tones with small numbers (Campbell & Scheepers, 2015; Oriet et al., 2005), it could be expected that participants' responses emphasize high-pitched vocalizations when the number is large and low-pitched vocalization when the number is small.

Finally, it has been found that in the parity judgement tasks, the participants often show longer reaction times for odd rather than even numbers (Hines, 1990). This phenomenon was explained by the so-called markedness theory, which proposes that most of the spatial and magnitude-related adjectives (e.g., long-short, far-near, high-low, etc.) are divided into pairs, one member being non-marked (e.g., long, far, high) and the other being marked (e.g., short, near, low). This proposal has been supported by the so-called MARC (Markedness Association of Response Codes) effect (Willmes & Iversen, 1995), in which the left-hand responses are associated with the odd numbers and the right-hand responses with the even numbers. Given that the left-hand can be considered as being linguistically marked concept, at least in right-handers (Huber et al., 2015), the MARC effect can be explained in terms of a congruity effect between the markedness of the number parity and the markedness of the label of the responding hand (Nuerk, Iversen, & Willmes, 2004). However, although the MARC hypothesis can be assumed to be generalizable to numerous spatial and magnitude-related concepts, to our knowledge, it has not been explored in relation to concepts other than the left and right hand. Consequently, given that long and high vocalizations have been proposed to be non-marked items of vocalization (Pulleyblank, 1983), we predicted that, in Experiment 1, long [ɑ]s would be produced faster in relation to even numbers, whereas short [ɑ]s are produced faster in relation to odd numbers. Similarly, in Experiment 2, we predicted that high-pitched [ɑ]s are produced faster in relation to even numbers, whereas low-pitched [ɑ]s are produced faster in relation to odd numbers as a result of linguistic markedness.

## Experiment 1

In this experiment, the participants were presented with the number 1, 2, 8, and 9. They were asked to judge the parity of the number by pronouncing the vowel [ɑ] in a short form or a long form. We measured reaction times of the onset of the vocalizations as well as their spectral components of intensity, $f_0$, F1 and F2. Predictions about how reaction times and the spectral components of intensity, $f_0$, and F1 could be influenced by the object size are discussed above. Regarding the F2, we did not have any strong pre-assumptions. However, given that F2 values largely comply with tongue fronting so that the more the tongue is pushed forward during vocalization, the higher is the F2 value (Fant, 1960), it is possible that the number size also somehow modulates these values. For example, given that large numbers have been associated with up-forward responses whereas small numbers are associated with down-backward responses (Hartmann, Gashaj, Stahnke, & Mast, 2014; Ito & Hatta, 2004), it is possible that large numbers similarly increase tongue fronting, which in turn could be observed in increased F2 values.

## Methods

### Participants

Twenty naïve volunteers participated in Experiment 1 (24–50 years of age; mean age = 29 years; 5 males). All participants were non-musicians, native speakers of Finnish and had normal or corrected-to-normal vision and were right-handed. We obtained written informed consent from all participants. The study was approved by the Ethical Review Board in Humanities and Social and Behavioural Sciences at the University of Helsinki.

### Apparatus, Stimuli, and Procedure

Each participant sat in a dimly lit room with his or her head 70 cm in front of a 19 in. CRT monitor (screen refresh rate 100 Hz; screen resolution 1280 × 1024). The head-mounted microphone was adjusted close to the participant mouth. The target stimuli consisted of four different centrally displayed numbers (1, 2, 8, and 9) that were written in Consolas font (black color; bold; font size: 100).

Each trial started with the presentation of a fixation cross (1° × 1°) for 800 ms. Then, the cross was replaced by an empty white screen, displayed for 700 ms. Next, the target stimulus appeared on the screen for 1000 ms. The participants were required to perform as fast and accurate vocal response as possible to the target. Reaction times were measured from the onset of the target object to the onset of the vocalization. The target stimuli were presented in random order with equal probability. Then, the target was replaced by an empty white screen for 1000 ms. All stimuli were presented on white background.

The participant was instructed to pronounce either short or long version of the vowel [ɑ] according to parity of the number. They were asked to use the level of intensity of their normal talking voice. They were not given any explicit instructions concerning the pitch of the voice. Half of the participants produced short [ɑ] if the number was odd and long [ɑ] if it was even (Mapping 1). The other half of the participants produced long [ɑ] if the number was odd and short [ɑ] if it was even (Mapping 2). Each participant was given as much practice as it took to perform the task fluently. In addition, the participant was allowed to have a break in the middle of the experiment. In total, the experiment consisted of 120 trials [30 × 4 (stimulus)].

The vocal responses were recorded for 2000 ms starting from the onset of the target object. At the beginning of the experiment, the recording levels were calibrated individually using the voice calibration function of the Presentation 16.1 software, so that the recording levels would match with the natural intensity of the participant's voice. In the calibration, the participants were required to pronounce the vowel [ɑ] approximately once every second. The calibration took around 1 min. Stimulus presentation and sound recording were done with the Presentation 16.1 software.

## Results

The vocal data were analyzed using Praat v. 5.3.49. Onsets and offsets of the vocalizations were first located individually for each trial. The intensity value was calculated as a peak value of the voiced section. The spectral components (F1 and F2) as well as $f_0$ were calculated as median values of the middle 1/3 of the voiced section.

### Reaction Times

Reaction times were measured from the onset of the target object to the onset of the vocalization. Errors and RTs more or less than two standard deviations from each participant's condition means were excluded from the reaction time analysis. Of the trials, 1.7% were removed as errors and 4.8% were removed as outliers. The combined removal of errors and outliers had left 93.5% of the raw data as correct responses. The condition means of these remaining data were computed for each participant and subjected to a repeated-measures analysis of variance (ANOVA) with the within-subjects variables of Number size (small [1&2] or large [8&9]) and Response (short [ɑ] or long [ɑ]), and the between-subjects variable of Mapping (Mapping 1 [short [ɑ]-odd/

long [ɑ]-even] or Mapping 2 [short [ɑ]-even/long [ɑ]-odd]). Although the same number of participants performed the experiment in both Mapping conditions, and hence that variable was counterbalanced, the Mapping was included to the design as the between-subjects because –as speculated in the Introduction– it is possible that we will observe a version of the MARC effect (Willmes & Iversen, 1995) in which the long [ɑ] is produced faster with even number and the short [ɑ] is produced faster with odd numbers. Post hoc comparisons were performed by means of t tests applying a Bonferroni correction when appropriate. A partial-eta-squared statistic served as an effect size estimate. Finally, for analyzing errors, all incorrect responses were converted to percentages and submitted to the ANOVA, similarly to the reaction time data.

The analysis of percentage of errors did not reveal any significant effects. This may be subject to ceiling effect given that the participants performed the task with high accuracy (98.3% correct responses). The analysis of reaction times revealed a main effect of Number size, F(1,18) = 8.67, MSE = 1787.13, p = 0.009, $\eta_p^2$ = 0.325. Responses were faster when the number was small (M = 584 ms; SEM = 21.4) rather than large (M = 594 ms; SEM = 21.8). More importantly, the analysis revealed a significant interaction between Number size and Response, F(1,18) = 31.19, MSE = 12730.67, $p < 0.001$, $\eta_p^2$ = 0.634. The short [ɑ] was produced faster when the number was small (M = 565 ms; SEM = 21.2) rather than large (M = 600 ms; SEM = 22.0) ($p < .001$). In contrast, the long [ɑ] was produced faster when the number was large (M = 588 ms; SEM = 22.1) rather than small (M = 604 ms; SEM = 22.4) (p = .010). This interaction is presented in Figure 1. The main effect of Mapping (p = .557) or the three-way interaction between Number size, Response and Mapping (p = .223) were not significant. Finally, when reaction times were analyzed including only the factor of Parity (odd or even), the main effect of Parity was not significant (p = .194). No other significant main effects or interactions were found.

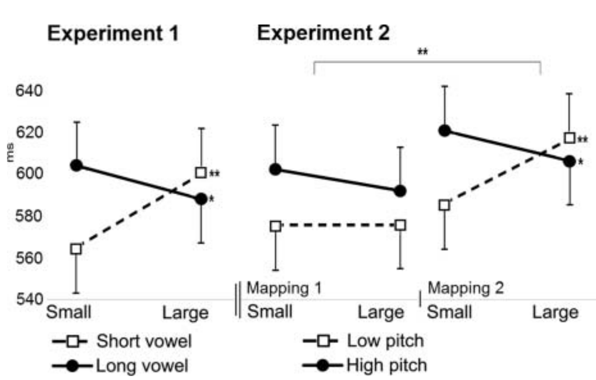


**FIGURE 1.** The mean vocal reaction times for Experiments 1 and 2 as a function of the number size (1 & 2 = small; 8 & 9 = large) and the type of the vocal response [Experiment 1: short vowel vs. long vowel; Experiment 2 (Mappings 1 and 2): Low pitch vowel vs. high pitch vowel]. Error bars depict the standard error of the mean. Asterisks indicate statistically significant differences (***p <. 001; **p < .01; *p < .05).

### Voice Characteristics

After removing the errors (1.7%) from the analysis of voice characteristics, the values over two standard deviations above or below each participant's condition means of vocalization length (2.1%), intensity (4.3%), $f_0$ (3.7%), F1 (5.5%), and F2 (4.9%) were also excluded from the voice characteristic analysis.

The analysis of vocalization length revealed a main effect of Response, F(1,18) = 234.04, MSE = 1233465.07, $p < 0.001$, $\eta_p^2$ = 0.929. As expected, the long vocalizations were longer (M = 378 ms; SEM = 18.6) than short vocalizations (M = 129 ms; SEM = 7.1). In addition, the analysis of intensity revealed a main effect of Response, F(1,18) = 11.28, MSE = 6.15, p = 0.003, $\eta_p^2$ = 0.385. Intensity was higher for short responses (M = 79.3 dB; SEM = 0.8) than for long responses (M = 78.8 dB; SEM = 0.8). The analysis of $f_0$ revealed a main effect of Response [F(1,18) = 11.59, MSE = 3179.77, p = 0.003, $\eta_p^2$ = 0.392] and Number size, F(1,18) = 12.02, MSE = 13.97, p = 0.003, $\eta_p^2$ = 0.400. $F_0$ was higher for short responses (M = 176.8 Hz; SEM = 11.4) than for long responses (M = 164.1 Hz; SEM = 9.6). In addition, $f_0$ was higher when Number size was large (M = 170.9 Hz; SEM = 10.4) than when it was small (M = 170.0 Hz; SEM = 10.3). Finally, regarding voice characteristics of F1 and F2, the only significant effect was a main effect of Response. Both values were higher in relation to short responses (F1: M = 605.3 Hz; SEM = 36.0; F2: M = 1132.0 Hz; SEM = 27.5) rather than long responses (F1: M = 561.4 Hz; SEM = 34.1; F2: M = 1080.6 Hz; SEM = 17.6) [F1: F(1,18) = 18.33, MSE = 38160.15, $p < 0.001$, $\eta_p^2$ = 0.505], [F2: F(1,18) = 11.46, MSE = 52471.09, p = 0.003, $\eta_p^2$ = 0.389]. No other significant main effects or interactions were found.

## Discussion

The results of Experiment 1 showed that participants produce the long [ɑ] faster when the number is large (8 and 9) rather than small (1 and 2), while the short [ɑ] is produced faster when the number is small rather than large. This finding can be assumed to replicate the effect reported by Kiesel and Vierck (2009), who found that short keypress responses are produced relatively rapidly when the viewed number is small, and long keypress responses are produced relatively rapidly when the number is large. As such, this study shows that magnitude information of a viewed number primes the processes responsible for selecting a vowel length for articulation.

The other significant effect (relevant for the current purposes) that was found in Experiment 1 was that $f_0$ of the vocalization was modulated by the size of the number. The vocal characteristic of $f_0$ was increased when the number was large in comparison to small. Although the effect was statistically significant ($p = .003$), the difference of $f_0$ related to small and large numbers was very small (0.9 Hz). A difference this small is hardly even perceptually noticeable, and therefore it is unlikely to reflect intentional attempts to produce high-low-pitched vocalizations according to the magnitude of the seen number. Instead, it rather reflects some involuntary and implicit articulatory modulations that are triggered by the number magnitude. The potential source of these modulations is discussed in detail in General Discussion.

Gentilucci et al., 2009 have similarly shown that when participants are presented with a graspable object while they have to pronounce a vowel [ɑ], $f_0$ component of the vocalization significantly increases by 1.4 Hz when the object is large in comparison to small. As such, the results of this study suggest that similar slight modulation in $f_0$ can be triggered not only by a size of the graspable object but also by magnitude information of a viewed number. However, given that the effect is very small, it has to be replicated in order to validate the results. In response to this requirement, Experiment 2 explores the effect further. In addition to replicating the $f_0$ effect, Experiment 2 aims at exploring whether the association between the number size and the pitch can be observed in reaction times when participants are required to pronounce the vowel [ɑ] in a low or high pitch according to the parity of the number. We assume that if it is indeed the case that large numbers are linked to high-pitched vocalization, this should be observed in heightened $f_0$ values and rapid high-pitched vocalizations (in comparison to low-pitched vocalizations) when the number is large. In contrast, small numbers should be linked to lower $f_0$ values and low-pitched vocalizations.

## Experiment 2

### Methods

*Participants*

Twenty naïve volunteers participated in Experiment 2 (22–42 years of age; mean age = 29 years; 6 males). All participants were non-musicians, native speakers of Finnish and had normal or corrected-to-normal vision. Two participants were left-handed. We obtained written informed consent from all participants. The study was approved by the Ethical Review Board in Humanities and Social and Behavioural Sciences at the University of Helsinki.

### Apparatus, Stimuli, and Procedure

The apparatus, stimuli, and voice calibration were the same as those in Experiment 1. The procedure was mostly similar to the one used in Experiment 1 with an exception that instead of pronouncing the vowel [ɑ] in the short and long forms, the participants were asked to pronounce the short [ɑ] in the low or high pitch according to the parity of the number. The participants were not given any specific instructions about how short the vocalization should be. They were only asked to refrain from producing long version of the vowel [ɑ]. The experiment was divided into two separate blocks, and there was a short break (approximately 5 minutes) between the blocks. In one block, the participants were required to pronounce low pitched [ɑ] if the number was even and high pitched [ɑ] if it was odd (Mapping 1). In another block, the participants were required to pronounce high pitched [ɑ] if the number was even and low pitched [ɑ] if it was odd (Mapping 2). In contrast to Experiment 1, the mapping condition was included to the design as a within-subjects variable in order to increase the statistical power of that variable. Indeed, it is possible that we did not observe any MARC effect between vowel length and the oddness/evenness of the number in Experiment 1 because the mapping condition was a between-subjects variable in the design. The order of the blocks was counterbalanced between the participants. Each participant was given as much practice, as it took to perform the task fluently. In total, the experiment consisted of 240 trials [30 × 4 (stimulus) x 2 (mapping)].

## Results

The vocal data were analyzed in the same way as in Experiment 1. Of the trials, 2.3% were removed as errors and 4.3% were removed as outliers. The combined removal of errors and outliers had left 93.5% of the raw data as correct responses. The condition means of these remaining data were computed for each participant and subjected to a repeated-measures analysis of variance (ANOVA) with the within-subjects variables of Mapping (Mapping 1: low-even/high-odd; Mapping 2: low-odd/high-even), Number size (small [1&2] or large [8&9]) and Response (low or high). Post hoc comparisons were performed by means of t tests applying a Bonferroni correction when appropriate. A partial-eta-squared statistic served as an effect size estimate. Finally, for analyzing errors, all incorrect responses were converted to percentages and submitted to the same ANOVA as the reaction time data.

### Reaction Times

The analysis of percentage of errors did not reveal any significant effects. This may be subject to ceiling effect given that the participants performed the task with high accuracy (97.7% correct responses). The analysis of reaction times revealed a main effect of Mapping [F(1,19) = 11.05, MSE = 28277.39, p = 0.004, $\eta_p^2 = 0.368$] and Response, F(1,19) = 5.85, MSE = 11102.26, p = 0.026, $\eta_p^2 = 0.235$. Responses were faster in Mapping 1 (M = 586 ms; SEM = 18.2) than in Mapping 2 (M = 612 ms; SEM = 17.7). In addition, responses were faster when the vowel was pronounced in low pitch (M = 591 ms; SEM = 16.0) rather than in high pitch (M = 608 ms; SEM = 19.5). Furthermore, the two-way interaction between Number size and Response was significant, F(1,19) = 9.02, MSE = 7203.23, p = 0.007, $\eta_p^2 = 0.322$. Responses were produced faster in high pitch when the number was large (M = 601 ms; SEM = 18.4) rather than small (M = 614 ms; SEM = 20.8) (p = .040). In contrast, responses were produced faster in low pitch when the number was small (M = 583 ms; SEM = 15.6) rather than large (M = 598 ms; SEM = 16.9) (p = .026). In addition, the three-way interaction between Mapping, Number size and Response was also significant, F(1,19) = 5.26, MSE = 2416.66, p = 0.033, $\eta_p^2 = 0.217$. As seen in Figure 1, the interaction between Number size and Response was only observed in Mapping 2. Finally, when reaction times were analyzed including only the factor of Parity (odd or even), the main effect of Parity was not significant (p = .244). The rest of the main effects and interaction were not significant.

### Voice Characteristics

After removing the errors (1.8%) from the analysis of voice characteristics, the values more or less than two standard deviations from each participant's condition means of vocalization length (1.8%), intensity (4.3%), $f_0$ (3.7%), F1 (5.5%), and F2 (4.9%) were also excluded from the voice characteristic analysis.

The analysis of vocalization length did not provide any significant main effects or interactions. The analysis of intensity revealed a main effect of response, F(1,19) = 19.62, MSE = 72.66, $p < 0.001$, $\eta_p^2 = 0.508$. Intensity was higher for high-pitched responses (M = 78.5 dB; SEM = 0.7) than for low-pitched responses (M = 77.1 dB; SEM = 0.8). The analysis of $f_0$ revealed a main effect of Response [F(1,19) = 96.71, MSE = 272861.99, $p < 0.001$, $\eta_p^2 = 0.836$]. $F_0$ was higher for high-pitched responses (M = 248.3 Hz; SEM = 16.4) than for low-pitched responses (M = 165.7 Hz; SEM = 10.3). More importantly, the analysis of $f_0$ also revealed a main effect of Number size [F(1,19) = 12.35, MSE = 14.17, p = 0.002, $\eta_p^2 = 0.394$]. However, the two-way interaction between Number size and Response [F(1,19) = 4.53, MSE = 13.16, p = 0.047, $\eta_p^2 = 0.192$] showed that the size of the number modulated responses only in relation to high-pitched responses (p = .003). In that condition, $f_0$ was higher when the number was large (M = 248.9 Hz; SEM = 16.4) rather than small (M = 247.7 Hz; SEM = 16.2). When the participants produced low-pitched responses, the effect was missing (p = .943). Regarding voice characteristics of F1 and F2, the only significant effect was a main effect of Response. Both values were higher in relation to high-pitched responses (F1: M = 460.0 Hz; SEM = 27.8; F2: M = 1136.5 Hz; SEM = 30.3) rather than low-pitched responses (F1: M = 378.9 Hz; SEM = 25.8; F2: M = 1104.7 Hz; SEM = 29.5) [F1: F(1,19) = 40.98, MSE = 263101.04, $p < 0.001$, $\eta_p^2 = 0.683$], [F2: F(1,19) = 4.66, MSE = 40293.41, p = 0.044, $\eta_p^2 = 0.197$]. No other significant main effects or interactions were found.

It can be seen in the Figure that the cases in which the participants are required to produce low-pitch responses to the number nine (i.e., Experiment 2, Mapping 2, Large stimulus, Low pitch) are produced relatively slowly, suggesting a strong mismatch between the number nine and the low pitch responses. This incongruency between the number nine and the low pitch responses can overemphasize the mapping effect (i.e., overall responses are produced significantly faster in Mapping 1 than in Mapping 2). Consequently, we reanalyzed the data by first removing all the reaction times that were associated with these conditions (number nine – low pitch), and then running ANOVA for the rest of the reaction times over all conditions including only the mapping as a factor. This analysis still showed a significant main effect for mapping, F(1,19) = 7,48, MSE = 5452.75, p = .013, $\eta_p^2 = .283$. The responses were still significantly faster in Mapping 1 (M = 586 ms; SEM = 18.1) than in Mapping 2 (M = 609 ms; SEM = 18.0). Consequently, is seems that although there appear to be emphasized incongruency between the low pitch responses and the number nine, which in turn emphasizes the mapping effect, this incongruency however is not the only component that produces the effect. The effect can be observed even when this component is entirely removed.

## Discussion

Experiment 2 replicated the $f_0$ effect that was observed in Experiment 1. However, the effect was only observed in high-pitched responses. When the participant performed high-pitched vocalizations, the $f_0$ of the vocalization was increased when the number was large in comparison to small. This $f_0$ effect was supported by the reaction time data showing that high-pitched vocalizations were produced relatively rapidly when the number was large and low-pitched vocalizations were produced relatively rapidly when the number was small. However, it is noteworthy that this reaction time effect was only observed in Mapping 2 in which high-pitched responses were performed to even numbers and low-pitched responses to the odd numbers. Below we further discuss these effects.

## General Discussion

This study provides the first investigation concerning interaction between vowel production and processes that represent the magnitude information of numbers. The study presents several novel findings related to processes underlying number cognition. Firstly, in Experiment 1, the participants displayed a tendency to associate the vocalization of the short vowel with small numbers and the vocalization of the long vowel with large number. This effect occurred even though the participants responded to the parity of the number and hence the magnitude information of the number was irrelevant to the task. As such, this finding is similar to the previously reported finding in which participants showed faster reaction times when they had to perform a keypress response for a short duration and the number was small, or when the keypress was performed for a long duration and the number was large (Kiesel, & Vierck, 2009). This study suggests that similar congruency effect between the number magnitude and response duration can be also observed in relation to vocal responses when the short and long responses have to be performed in the context of short and long vowels. If our finding is considered in the context of the ATOM hypothesis (Walsh, 2003), it can be proposed that this study supports the view according to which there is some level of overlap in sensory-motor processes that encode magnitude information for numbers and duration. Consequently, the processes involved in planning the time dimension on articulatory planning processes are automatically influenced by the concurrently processed information about number magnitude, which in turn biases response selection processes related to producing a short or long vowel. Due to these associations between representing duration information for vocalization and magnitude, it is not surprising that people tend to vocally emphasize the magnitude of a given attribute by stretching a vowel (e.g., “the weather was *sooo* cold”).

The results of Experiment 2 revealed that participants prefer to produce high-pitched vocalizations to odd numbers and low-pitched vocalizations to even numbers rather than the other way around. This effect can be assumed to be a parallel cognitive phenomenon to the MARC effect (Willmes & Iversen, 1995), in which the left hand responses are performed faster to odd numbers and the right hand responses are performed faster to even numbers. These effects have been generally explained by linguistic markedness account (Zimmer, 1964). As mentioned in Introduction, this account deals with conceptual pairs of complementary adjectives such as high-low, good-bad and odd-even. The non-marked adjective is referred to be the more natural or basic form of the concept pair. For example, even numbers are assumed to be non-marked items of the concept pair “odd-even,” because when learning basic multiplication, the correct answer is an even number at 0.75 probability—thus making even numbers more familiar than odd numbers (Lochy, Seron, Delazer, & Butterworth, 2000). Correspondingly, at least for the right-handers (Huber et al., 2015), the right hand is more commonly used than the left hand, which makes the right hand the non-marked basic item of the left-right concept pair. It is commonly assumed that the MARC effect is caused by the congruency between the markedness of the number parity and the markedness of the label of the responding hand (Nuerk et al., 2004). From this perspective, it can be proposed that the participants preferred to produce high-pitched vocalization with odd numbers and low-pitched vocalization with even numbers because there is decreased congruency in markedness dimensions between the stimulus and response when responses are performed in a reversed mapping condition. However, it has to be noticed that the odd effect (Hines, 1990), in which participants typically respond slower to odd number than to even numbers, was not observed in this study. This notion might be considered to dilute the validity of our proposed explanation to the mapping effect observed in Experiment 2. However, cognitive mechanisms underlying the odd effect and the MARC effect are not understood that clearly that one could undoubtedly state that observing the odd effect is obligatory condition for observing the MARC effect. Therefore, it is possible that the processes responsible for mapping marked/unmarked responses to the odd/even numbers can produce the MARC effect even when the processing of odd and even numbers would not be biased enough to produce a noticeable odd effect.

Contrary to the current finding, research in phonetics has proposed that high tone is in fact non-marked with respect to low tone (e.g., Pulleyblank, 1983). According to this view, the high-pitched vocalization in our data should have been associated with the even numbers, and the low-pitched vocalization with the odd numbers. However, it has to be highlighted that markedness of the opposing items of the concept pair is relative rather than absolute issue. Given that, in phonetics, the marked phonetic structures are less natural, more effortful (i.e., harder to articulate), less common, less expected and perceptually more salient (De Lacy, 2007), it would be plausible to assume that at least in the present experimental set up, the high-pitched vocalization was more likely to be marked than the low-pitched vocalization. That is, because the vocalizations were produced significantly faster in low pitch than in high pitch, thus supporting the view that the low-pitched vocalization required less articulatory effort than producing the high-pitched vocalization. In addition, it has been noticed that pitch accents in intonation are usually high on “new” information on the discourse (Bolinger, 1986) underlying unexpectedness of this information. In addition, raising a pitch is commonly used prosodic cue when one is, for example, indicating a question or when one wants to underline a specific word in a sentence (Ohala, 1983). In this sense, by raising the pitch in relation to the average pitch of a given sentence one is able to mark a specific component of the sentence so that it stands out from the stream of speech and directs attention of another person to that component. As such, we propose that in the context of vowel production in general, and in this study in particular, the high-pitched vocalizations were marked with respect to low-pitched vocalizations. Consequently, the participants showed tendency to pair low-pitched vocalizations with even numbers and high-pitched vocalizations with odd numbers. This effect suggesting congruency between high-odd and low-even pairs appears to be particularly emphasized by a strong incongruency between the number nine and low pitched vocalizations that can be seen in particularly slow reaction times in Figure 1 (Experiment 2; Mapping 2) when the number is large and the vocalization response is low.

Interestingly, in addition to the linguistic markedness explanation, the odd effect and the MARC effect have been also explained by the polarity correspondence account (Proctor & Cho, 2006). According to that account, these parity effects reflect associating "even" and "right-hand" with positive polarity and "odd" and "left-hand" with negative polarity. This view assumes that congruency in polarity between the stimulus and response should lead to facilitated responses. Because concepts such as long and high similarly associate with positive polarity, they should also associate with even numbers. However, in Experiment 1, the participants' short-long responses were not significantly influenced by the number parity, while in Experiment 2, the results showed a pattern that is better in line with the markedness account than with the polarity correspondence account. Therefore, we propose that the polarity correspondence account of the parity effects was not supported by the results of this study. However, this argument should be treated with caution because one possible reason for the observed MARC effect in Experiment 2 but not in Experiment 1 is that in Experiment 2 the mapping condition was included to the design as a within-subjects variable, while in Experiment 1 it was included as a between-subjects variable that can be assumed to decrease its statistical power.

This study also showed that high-pitched vocalizations were performed faster when the number was large rather than small, whereas low-pitched vocalizations were performed faster when the number was small rather than large. However, this effect was only evident in Mapping 2 when the participants produced a high-pitched vocalization to odd numbers and a low-pitched vocalization to even numbers. Hence, it seems that the previously mentioned markedness congruency effect between the required pitch of the response and the parity of the number confounds with the congruency effect between the required pitch of the response and the size of the number. One potential reason for this is that congruency effects between the size of the number and the type of response such as the SNARC effect is likely to become larger with slower responses (Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006). The magnitude information of the number presented to the participant has no time to modulate response planning processes when responses are performed too rapidly. Therefore, it is possible that in Mapping 1, in which there was a strong congruency between the required pitch of the response and the parity of the number, the responses were performed particularly smoothly and rapidly. Hence, there was no sufficient time for the effect between the required pitch of the response and the size of the number to build up under the influence of number magnitude information.

The study also showed that the number size can automatically modulate $f_0$ component of the voice. $F_0$ was higher with large numbers in comparison to small numbers, thus complying with the previous findings that people associate high pitch sounds with large numbers and low pitch sounds with small numbers (Oriet et al., 2005). This outcome was also predicted based on the previously found effect showing increase in $f_0$ when the participants are presented with large objects in comparison to small objects (Gentilucci et al., 2009). This study shows that not only are large graspable objects associated with relatively high-pitched vocalizations, but large numbers are also associated with relatively high-pitched vocalizations. In addition, this finding is in line with the above mentioned congruency effect between the required pitch of the response and the size of the number. That is, it appears that large numbers do not only speed up producing high-pitched vocalizations in comparison to low-pitched vocalizations, but they also slightly increase the values directly linked to the heightening of vocal pitch. Although this $f_0$ effect was observed in both experiments of the study, in Experiment 2 the effect was only observed in relation to the high-pitched vocalizations. This suggests that large numbers increase the $f_0$ values when participants are required to perform the vowel in their normal pitch (Experiment 1) or when they have to perform atypically high-pitched vocalization (Experiment 2). Perhaps the $f_0$ effect reflects the influence of relatively large numbers on processes that are responsible for heightening a pitch rather than the influence of relatively small numbers on processes responsible for lowering a pitch. Consequently, when the participants are required to produce atypically low-pitched vocalizations, the number magnitude does not influence the vocal processes related to pitch production.

Why then are large numbers associated with heightened pitch of a voice? It has previously been shown that when participants are required to perform vertically aligned manual responses to auditory sounds, upper responses are performed relatively rapidly when the pitch of the sound is high, whereas lower responses are performed relatively rapidly when the pitch is low (Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006). In addition, vertically aligned manual responses are also performed faster with upper responses when the number is large and with the lower responses when the number is small (Hartmann et al., 2014; Ito & Hatta, 2004). These findings have been explained, for example, by the polarity correspondence account (Cho & Proctor, 2007). According to this account, the high response, large number and high sound are all positive polarity concepts, whereas the low response, small number and low sound are all negative polarity concepts. Responses are facilitated when polarity dimensions of the stimulus and response are congruent. Similarly, the present findings can be explained by this polarity account so that the high-pitched vocalizations are preferably mapped with large numbers and the low-pitched vocalizations with small numbers because of the congruency in the polarity dimensions.

The polarity account explains the reaction time effect between the required pitch of response and the number size in relatively comprehensive manner. The polarity account is theoretically suitable for dealing with spatial stimulus-response compatibility effects and occur when the stimulus and response alternatives are coded categorically in relation to positive and negative polarities (Cho, Bae, & Proctor, 2012). However, it is more difficultly employed explaining the effect showing an automatic heightening of $f_0$ triggered by the large numbers. This effect is more likely to be based on some basic sensory-motor processes that implicitly anchor conceptual magnitude information to motor representations. However, in order to clarify our view concerning why and how the number magnitude is associated with pitch in the way proposed by the current findings, it is important to understand why and how high and low pitch sounds could be grounded on vertical body movements. Relevantly for our view, Bolinger (1983; 1986) has suggested that pitch changes in intonation are processed in integration with bodily gestures: pitch and body parts move up and down together providing audible and visible prosody, respectively. Indeed, it has been shown that spontaneous head lowering is associated with intonationally falling pitch and head rising with rising pitch (McClave, 1991). Similar phenomenon has been also recognized in singing. The head extension is a commonly used method of singers (Miller, 2000) that might assist laryngeal positioning to ease the production of high notes (see Knight, 2013). Indeed, people have a tendency to raise their head in relation to high-pitched sounds and lower their head in relation to low-pitched sounds (Horstmann & Ansorge, 2011). Moreover, it has been also shown that the $f_0$ component of vocal spectra increases when the head is raised in comparison to lowering the head (Knight, 2013). Taken together, people show a tendency to associate head raise with high sounds and head lowering with low sounds. Ultimately, this tendency might arise from speech production mechanisms so that these vertical head movements assist laryngeal positioning for producing high and low-pitched vocalizations. Finally, in evolution of speech, this tendency might have been adapted to provide visual boosting for intonational auditory cues.

Why then would the sensory-motor processes map a high pitch with large magnitudes and a low pitch with small magnitudes? Firstly, our view assumes that when a participant is representing the magnitude of a number at the conceptual level, the sensory-motor system maps large numbers onto a mental representation of the upper space because the ground represents the natural zero-point for vertically increasing magnitude (Clark, 1973; Fischer, & Brugger, 2011). Because this magnitude information is automatically encoded in relation to corresponding motor processes (Bueti, & Walsh, 2009; Walsh, 2003), viewing large numbers excite upward-directed body movements of the eyes (Schwarz, & Keus, 2004) and hands (Hartmann et al., 2014; Ito, & Hatta, 2004), for example. The same phenomena might also occur in relation to head movements. It is possible that the $f_0$ was influenced by the number magnitude in this study because large numbers implicitly triggered a slight head raise that modulated the laryngeal positioning, which in turn heightened the $f_0$ values of the vocalization. Consequently, rather than explaining the interaction between the number size and the pitch-related aspects of the response by the polarity account, concordantly to the ATOM hypothesis (Walsh, 2003), we prefer the number embodiment account in our explanation. According to this view, large numbers are connected to high-pitched sounds in the shared representational medium that processes magnitude information and pitch highness in relation to the processes that are responsible for planning vertical body movements.

In conclusion, this study shows for the first time that number magnitude is partially represented in integration with vocalization processes. The production of relatively long and high-pitched vowel was associated with large numbers and short and low-pitched vowel was associated with small numbers. In addition, it was found that high-pitched vocalizations are preferably mapped with the odd numbers, while the low-pitched vocalizations are mapped with the even numbers. These findings support the view that motor processes contribute to representing magnitude information in general and number magnitude in particular (Walsh, 2003). In addition to emphasizing involvement of manual motor processes in representing magnitude information (e.g., Andres et al., 2004), this study suggests that relevance of articulatory motor processes should be also emphasized when investigating processes underlying the representation of magnitude information.

## ACKNOWLEDGMENT AND FUNDING DETAILS

The research was supported by the Academy of Finland (grant number 1265610). We would like to thank Liisa Salmela, Anna-Maria Hirvonen and Katja Putkinen for helping in collecting data for Experiment 1.

## ORCID

*M. Vainio* http://orcid.org/0000-0003-2570-0196

## REFERENCES

Alards-Tomalin, D., Walker, A. C., Shaw, J. D., & Leboe-McGowan, L. C. (2015). Is 9 louder than 1? Audiovisual cross-modal interactions between number magnitude and judged sound loudness. *Acta psychologica*, *160*, 95–103. doi:10.1016/j.actpsy.2015.07.004

Alards-Tomalin, D., Walker, A. C., Nepon, H., & Leboe-McGowan, L. C. (2017). Dual-task interference effects on cross-modal numerical order and sound intensity judgments: The more the louder? *The Quarterly Journal of Experimental Psychology*, *70*(9), 1943–1963. doi:10.1080/17470218.2016.1216139

Andres, M., Davare, M., Pesenti, M., Olivier, E., & Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport*, *15*, 2773–2777.

Andres, M., Ostry, D. J., Nicol, F., & Paus, T. (2008). Time course of number magnitude interference during grasping. *Cortex*, *44*, 414–419. doi:10.1016/j.cortex.2007.08.007

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28*, 105–124. doi:10.1017/S0140525X05000038

Bolinger, D. (1983). Intonation and gesture. *American Speech*, *58*, 156–174. doi:10.2307/455326

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*, 1831–1840. doi:10.1098/rstb.2009.0028

Campbell, J. I., & Scheepers, F. (2015). Effects of pitch on auditory number comparisons. *Psychological research*, *79*(3), 389–400. doi:10.1007/s00426-014-0571-7

Cho, Y. S., & Proctor, R. W. (2007). When is an odd number not odd? Influence of task rule on the MARC effect for numeric classification. *Journal of Experimental psychology: Learning, memory and Cognition*, *33*, 832.

Cho, Y. S., Bae, G. Y., & Proctor, R. W. (2012). Referential coding contributes to the horizontal SMARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 726. doi:10.1037/a0026157

Clark, H. H. (1973). Space, time, semantics, and the child. *Cognitive development and the acquisition of language*, *27*, 63.

Coward, S. W., & Stevens, C. J. (2004). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *The Psychological Record*, *54*, 349. doi:10.1007/BF03395478

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371. doi:10.1037/0096-3445.122.3.371

de Lacy, P. (2007). The interaction of tone, sonority, and prosodic structure. In P. de Lacy (Eds.), *The Cambridge handbook of phonology* (pp. 281–307). Cambridge: Cambridge University Press.

Ellis, R., Tucker, M., Symes, E., & Vainio, L. (2007). Does selecting one visual object from several require inhibition of the actions associated with nonselected objects? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 670. doi:10.1037/0096-1523.33.3.670

Fant, G. (1960). *Acoustic theory of speech perception*. The Hague: Mouton.

Fischer, M. (2003). Spatial representations in number processing–evidence from a pointing task. *Visual cognition*, *10*, 493–508. doi:10.1037/rev0000038

Fischer, M. H., & Brugger, P. (2011). When digits help digits: Spatial–numerical associations point to finger counting as prime example of embodied cognition. *Frontiers in Psychology*, *2*, 1–7. doi:10.3389/fpsyg.2011.00260

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, *68*, 1191–1203. doi:10.3758/BF03193720

Gentilucci, M., Benuzzi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology*, *86*, 1685–1699. doi:10.1152/jn.2001.86.4.1685

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*, 949–960. doi:10.1016/j.neubiorev.2006.02.004

Gentilucci, M., Campione, G. C., Dalla Volta, R., & Bernardis, P. (2009). The observation of manual grasp actions affects the control of speech: A combined behavioral and transcranial magnetic stimulation study. *Neuropsychologia*, *47*, 3190–3202. doi:10.1016/j.neuropsychologia.2009.07.020

Gentilucci, M., & Campione, G. C. (2011). Do postures of distal effectors affect the control of actions of other distal effectors? Evidence for a system of interactions between hand and mouth. *PLoS One*, *6*(5), e19793. doi:10.1371/journal.pone.0019793

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 32. doi:10.1037/0096-1523.32.1.32

Hartmann, M., Gashaj, V., Stahnke, A., & Mast, F. W. (2014). There is more than "more is up": Hand and foot responses reverse the vertical association of number magnitudes. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1401. doi:10.1037/a0036686

Hines, T. M. (1990). An odd effect: Lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*, 40–46. doi:10.3758/BF03202644

Horstmann, G., & Ansorge, U. (2011). Compatibility between tones, head movements, and facial expressions. *Emotion*, *11*, 975. doi:10.1037/a0023468

Huber, S., Klein, E., Graf, M., Nuerk, H. C., Moeller, K., & Willmes, K. (2015). Embodied markedness of parity? Examining handedness effects on parity judgments. *Psychological Research*, *79*, 963–977. doi:10.1007/s00426-014-0626-9

Ito, Y., & Hatta, T. (2004). Spatial structure of quantitative representation of numbers: Evidence from the SNARC effect. *Memory & Cognition*, *32*, 662–673. doi:10.3758/BF03195857

Kiesel, A., & Vierck, E. (2009). SNARC-like congruency based on number magnitude and response duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 275. doi:10.1037/a0013737

Kim, A., & Zaidel, E. (2003). Plasticity in the SNARC effect during manipulation of order in response conditions. *Journal of Cognitive Neuroscience*, Supplement, 134.

Knight, E. J. (2013). *The effect of head flexion/extension on acoustic measures of singing voice quality* (Doctoral dissertation). University of North Texas.

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Boston, MA: Thomson Wadsworth.

Lindemann, O., Abolafia, J. M., Girardi, G., & Bekkering, H. (2007). Getting a grip on numbers: Numerical magnitude priming in object grasping. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1400. doi:10.1037/0096-1523.33.6.1400

Lochy, A., Seron, X., Delazer, M., & Butterworth, B. (2000). The odd-even effect in multiplication: Parity rule or familiarity with even numbers? *Memory & cognition*, *28*, 358–365. doi:10.3758/BF03198551

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*, 873–881. doi:10.1177/0956797610370158

McClave, E. (1991). *Intonation and gesture*. Washington DC: Doctoral Dissertation, Georgetown University.

Miller, R. (2000). *Training Soprano Voices*. Oxford: Oxford University Press.

Moretto, G., & Di Pellegrino, G. (2008). Grasping numbers. *Experimental Brain Research*, *188*, 505–515. doi:10.1007/s00221-008-1386-9

Nuerk, H. C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Quarterly Journal of Experimental Psychology Section A*, *57*, 835–863. doi:10.1080/02724980343000512

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech* (pp. 189–216). New York: Springer.

Oriet, C., Tombu, M., & Jolicoeur, P. (2005). Symbolic distance affects two processing loci in the number comparison task. *Memory & Cognition*, *33*, 913–926. doi:10.3758/BF03193085

Parise, C. V., & Spence, C. (2009). 'When birds of a feather flock together': Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One*, *4*(5), e5664. doi:10.1371/journal.pone.0005664

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*, 416. doi:10.1037/0033-2909.132.3.416

Pulleyblank, D. (1983). Tone in lexical phonology. PhD dissertation, MIT.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*, 188–194. doi:10.1016/S0166-2236(98)01260-0

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, *99*, 113–129. doi:10.1016/j.cognition.2005.01.004

Srinivasan, M., & Carey, S. (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*, 217–241. doi:10.1016/j.cognition.2010.05.005

Schwarz, W., & Keus, I. M. (2004). Moving the eyes along the mental number line: Comparing SNARC effects with saccadic and manual responses. *Attention, Perception, & Psychophysics*, *66*, 651–664. doi:10.3758/BF03194909

Tiainen, M., Tiippana, K., Vainio, M., Peromaa, T., Komeilipoor, N., & Vainio, L. (2016). Selective influences of precision and power grips on speech categorization. *PloS one*, *11*(3), e0151688. doi:10.1371/journal.pone.0151688

Vainio, L., Schulman, M., Tiippana, K., & Vainio, M. (2013). Effect of syllable articulation on precision and power grip performance. *PloS one*, *8*(1), e53061. doi:10.1371/journal.pone.0053061

Vainio, L., Rantala, A., Tiainen, M., Tiippana, K., Komeilipoor, N., & Vainio, M. (2017). Systematic influence of perceived grasp shape on speech production. *PloS one*, *12*(1), e0170221. doi:10.1371/journal.pone.0170221

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488. doi:10.1016/j.tics.2003.09.002

Willmes, K., & Iversen, W. (1995 April). On the internal representation of number parity. The conference paper presented at the *Spring Annual General Meeting of the British Neuropsychological Society*. London.

Zimmer, K. (1964). Affixed negation in English and other languages: An investigation of restricted productivity. *Word*, *20*, 2, Monograph No. 5.

*Received August 11, 2017*
*Revised January 23, 2018*
*Accepted January 27, 2018*