Journal of Pacific Rim Psychology

www.cambridge.org/prp

Effects of syllable name change on pitch comparison

Zhiqiao Wang and Qingfen Hu

Beijing Normal University, Beijing, China

Original Article

Cite this article: Wang Z. and Hu Q. (2020) Effects of syllable name change on pitch comparison. *Journal of Pacific Rim Psychology*, Volume 14, e1. https://doi.org/10.1017/prp.2019.25

Received: 25 March 2019
Revised: 7 November 2019
Accepted: 10 November 2019

Keywords:
syllable name; pitch comparison; music; sequential representation; magnitude representation; spatial representation

Author for correspondence:
Qingfen Hu, Email: qingfenhu@bnu.edu.cn

Abstract

Previous research has found that musicians' pitch judgments, unlike non-musicians', are influenced by syllable names. Although non-musicians fail to identify absolute pitches, they acknowledge the direction of pitch change. The present experiment investigated whether non-musicians' judgments of pitch change can be influenced by the direction of syllable name change. Moreover, we examined the spatial, magnitudinal and sequential nature of pitches and syllable names. Participants ($N = 33$) were asked to hear two successive tones sung by syllable name and to judge the direction of pitch change by pressing vertically arranged buttons. Participants' accuracy of pitch change judging was found to be influenced by the direction of syllable name change. However, the response location was not found to interact with pitch change or syllable name change. The distance effect was found in pitches but not in syllable names. A sequence effect was found that trials with early-in-sequence syllable names were responded faster than trials with late-in-sequence syllable names. These results suggest that syllable names can influence non-musicians' pitch judgments in a relative context. We suggest that it is the sequential order of syllable names that is the product of cultural activities that interfere with the judgment of pitch change.

Music is a human-specific sound system that conveys complicated information and emotion (Patel, 2012). The music system divides the continuum of physical vibration frequencies into discrete units (i.e., tones) and labels them under certain rules. In Western music theory, which is widely used across the world, musicians define an octave as the distance between two tones, the ratio of whose frequencies is 2. Within an octave, solmization labels the seven discrete tones with seven sequential syllable names, which are usually referred to as dol, ray, mi, fa, sol, la, and xi.

Syllable names are often used when singing musical scales or melodies. Compared to music played by instruments, *tones sung with syllable names* possess both pitches (i.e., the frequency of the music sound) and syllable names (i.e., a monosyllable). The relationship between the syllable names and pitches sung within a tone can be either fixed (i.e., fixed dol-solmization) or dynamic (i.e., movable dol-solmization).

Although neither solmization system dominates the other, syllable names do seem to have a strong, fixed relationship with pitches among absolute pitch possessors (AP possessors) and well-trained musicians, in which one single syllable name corresponds to one single pitch. Itoh, Suwazono, Arao, Miyazaki, and Nakada (2005) provided evidence for the existence of this fixed syllable name-pitch relationship using an auditory Stroop paradigm. They showed nine auditory stimuli (i.e., C3, D3, and E3 sung with dol, ray, and mi), in which the syllable name either corresponds to the pitch or does not, to AP possessors and asked them to report the tone's pitch. AP possessors responded more rapidly to corresponding tones than non-corresponding tones. This result indicated that the syllable names activated specific pitch representations automatically, which interfered with the process of pitch identification. Similar effects were observed in multiple tasks and in the visual modality. However, they were not found among non-musicians (Akiva-Kabiri & Henik, 2012; Grégoire, Perruchet, & Poulin-Charronnat, 2013; Morgan & Brandt, 1989; Schulze, Mueller, & Koelsch, 2013). Negative results of non-musicians suggested that there is not such a fixed relationship connecting syllable names and pitches so closely in non-musicians.

Although non-musicians cannot identify absolute pitch, they can identify the relative heights of pitches through comparison (i.e., relative pitch). Non-musicians are good at identifying the direction of pitch change (ascending or descending) although they cannot recognize the distance between them (McDermott & Oxenham, 2008). Similar to pitches, syllable names are also arranged sequentially within an octave. That is, ray only appears after dol, and mi only appears after ray; the later the syllable name appears in the sequence, the higher the pitch it represents. At the same time, syllable names covary with musical pitches on many occasions and in a certain range (i.e., an octave). For instance, in fixed dol-solmization, syllable names correspond with pitches one-to-one; in movable dol-solmization, although the relation between syllable names and pitches are changeable, the covariation remains under a certain key.

© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE UNIVERSITY PRESS

Then, do non-musicians also represent the syllable names in a relative manner – that is, a monotonous sequence – and are they sensitive to this order? Do non-musicians, who cannot identify the fixed correspondence between absolute pitches and their corresponding syllable names, recognize the directions of both the syllable name change and the pitch change? If they do, will these two directions be associated and interfere with each other? This is an interesting and undiscussed issue.

Another issue worth exploring is which kind of representation is involved in the interference between pitch and syllable name change. The representation of pitch change has a spatial component. Studies have shown that when non-musicians were asked to compare two successive pure tones (pitch comparison task) and react via two vertically arranged buttons on a keyboard, they responded to higher tones faster and more accurately with the far button than with the near one and to lower tones faster and more accurately with the near button than with the far one (i.e., the Spatial-Musical Association of Response Codes effect, or SMARC effect; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). The SMARC effect indicates the spatial nature of pitch height representation: a higher pitch is represented in a higher place, whereas a lower pitch is represented in a lower place in mental space. This spatial representation can be activated automatically and interferes with the judgment of other stimuli that also have spatial representations, such as numbers (Campbell & Scheepers, 2015).

There is little evidence on the characteristics of the representation of syllable names. As symbolic labels of musical pitches, syllable names appear in a monotonous manner and covary with musical pitches, as mentioned earlier. Three kinds of representations might be activated due to these features of syllable names. First, it is possible that syllable names also activate a spatial representation like pitches do. If this is true, there would be a SMARC-like interference between syllable name change and the reaction places.

Second, it is possible that the syllable names activate magnitude representation like numbers and many other quantitative sequences. The identifying feature of magnitude comparison is the distance effect. That is, the response time to make a comparison between two close magnitudes is longer than the comparison between two distant magnitudes (Campbell & Sacher, 2012). The distance effect can also be found in pitch comparison tasks in non-musicians (Kadosh, Brodsky, Levin, & Henik, 2008), indicating that non-musicians' representation of pitch is also magnitude. If syllable names activate magnitude representations, there would be a distance effect in the comparison of two syllable names; the response time of the comparison between two close syllable names would be longer than that of two distant ones. In this way, the interference between syllable name changes and pitch changes might be in terms of magnitude rather than spatial.

The last possibility is that syllable names only activate a sequential representation. Before representing the syllable names, one mentally searches the sequence one by one, and terminates the searching process once the target is found. As a result, the reaction time (RT) in tasks involving sequential representation depends on the position the target units are located (Kosslyn, Ball, & Reiser, 1978). Halpern (1988) examined how people operate on representations of familiar songs. Participants were shown a lyric from the beginning part of a familiar song, then shown another one (the target lyric). They were asked to decide whether the target was from the same song as the first. Results showed that the later the target appeared in the familiar song, the longer time it took participants to react. Similarly, if syllable names are represented sequentially, we will observe a sequence effect, in which the reaction to trials with early syllable names (i.e., the earlier ones in the sequence that are close to dol) takes longer time than those with late syllable names (i.e., the late ones in the sequence which are far from dol).

The present study examined the possible existence of interference between the syllable name change and the pitch change, as well as the possible representations of these two dimensions. In the experiment, participants listened to two auditory stimuli. The task was to determine whether the pitch of the target sound (the second auditory stimulus) was higher or lower than that of the anchor sound (the first auditory stimulus) and to react with two vertically arranged mouses. The auditory stimuli were tones sung with syllable names. Therefore, both syllable names and pitches changed across the two sounds. In half of the trials, the direction of syllable name change was congruent with that of the pitch change (i.e., higher pitch-higher syllable name or lower pitch-lower syllable name), while in the other half of the trials, the direction of syllable name change was incongruent with that of pitch change (i.e., higher pitch-lower syllable name or lower pitch-higher syllable name). If the two syllable names automatically activated a mental representation in which they are directionally arranged according to the pitch heights they represent, there would be interference between these two dimensions. That is, the congruent trials would be responded to faster and more accurately than the incongruent trials.

Then, we tested the possible characteristics of representations for both pitches and syllable names. First, in the present study, half of the participants were assigned to the accordant task mapping condition (i.e., higher pitch-far place or lower pitch-near place), while the other half were assigned to the reversed task mapping condition. If the pitches are represented spatially, being associated with vertical space (i.e., activating a spatial representation), we would observe a task effect that participants in the accordant task mapping group outperformed those in the reversed task mapping group (the SMARC effect). Similarly, we analyzed the interaction between syllable name change and react-place to investigate the possible spatial representation of syllable name change. If syllable names are associated with vertical space, we would observe better performance in trials with ascending syllable names compared to descending ones in far places and in trials with descending syllable names compared to ascending ones in near places (the SMARC-like effect).

Second, we tested the distance effect in the comparison. In the present study, half of the target sounds were at least 5 semitones from the anchor, and the other half were 2 semitones at most. If there is a distance effect in pitch height, the speed of comparison through longer pitch distance will be faster than that through shorter pitch distance. Similarly, the syllable names of the target sounds were also set to be close or far from the anchor. If there is a distance effect in syllable names, the target sounds with syllable names far from that of the anchor sound will have shorter response times.

At last, our experimental design gave us a chance to examine the sequential effect if the examination of the SMARC-like effect and the distance effect failed to provide positive evidence. For example, there were four kinds of possible syllable names in the target sounds: dol, mi, sol, and xi; dol and mi appear early in the sequence of syllable name sequence while sol and xi appear in the later half. If syllable names activate sequential representation, we would observe shorter reaction time in early syllable name trials (whose

https://doi.org/10.1017/prp.2019.25 Published online by Cambridge University Press

target sounds' syllable names were dol and mi) than late syllable name trials (those whose target sounds' syllable names were sol and xi).

To sum up, our main hypothesis was that the change in syllable names would interfere with the change in pitch among non-musicians in a pitch comparison task. We would expect better task performance to occur when syllable name changes are congruent with pitch changes. Moreover, we would expect the SMARC effect, in which pitch change activates a spatial representation to occur, resulting in better task performance when ascending/descending pitches are mapped in far/near places respectively (accordant task mapping), than reversed (reversed task mapping). We were also keen to see whether a SMARC-like effect would occur between syllable name change and react-place, resulting in better task performance when ascending/descending syllable names are reacted to in far/near places respectively than reversed, which indicates the spatial representation of the syllable name changes. Furthermore, we expected the distance effect for both pitch and syllable name changes, which would indicate the magnitude representations of these two changes. At last, if the SMARC-like effect and the distance effect were absent, we would examine the sequential effect, which would indicate the sequential representation of monotonously arranged units.

Methods and materials

Participants

The participants were 33 undergraduate and postgraduate students (4 males and 29 females) recruited from Beijing Normal University as paid volunteers. The average age of the participants was 21.03 years, all right-handed. Thirteen extra participants were excluded from the final sample, in six cases because they did not change hand postures, one because she used unexpected strategy (singing along with audio stimuli), and six due to low accuracy (less than 80%). In the final sample, 18 participants (4 males, 14 females) were left in the accordant task mapping group and 15 (all female) were in the reversed task mapping group. No participant in the final sample had received singing or instrumental training for more than three years. The protocol of study was approved by Institutional Review Board of the Faculty of Psychology, Beijing Normal University. Written informed consent was provided by each participant. G*Power analysis indicated that at least 34 subjects were required to detect effects with an effect size $f = .25$, alpha = .05, and power = .80 when the type of test was a mixed-design ANOVA.

Materials

Each trial consisted of one anchor sound and one target sound. They were made using NIAONIAO Virtual Singer software (ver2.2.00) and further edited using GoldWave software (ver5.67). In each sound slice, a female human voice sang a syllable name in a certain pitch. For anchor sounds, only one piece of sound slice was used, namely 'fa' in the pitch of F4. For target sounds, 16 pieces of sounds with different pitches and syllable names were used. There were four levels of pitch height: C4, E4, G4 and B4. In half of the target sounds, the pitches (G4 and B4) were higher than the anchor sound (ascending pitch change), one (B4) of which was far from the anchor and the other (G4) of which was near. In the other half of the target sounds, the pitches (C4 and E4) were lower than the anchor sound (descending pitch change); also, one (C4) was far from the anchor and the other (E4) was near. Four syllable names were used in the target sounds: dol, mi, sol, and xi, in which dol and mi were 'earlier' than the anchor (descending syllable name change) and sol and xi were 'later' than the anchor (ascending syllable name change). Similarly, dol and xi were far from the anchor's syllable name, while mi and sol were near. Table 1 shows all 16 combinations of pitches and syllable names used in the target sounds. Therefore, in half of the trials, pitch change (anchor-target relation of pitch) was congruent with syllable name change (anchor-target relation of syllable names), whereas in the other half of the trials, pitch change was incongruent with syllable name change. Each trial was repeated four times in a block, resulting in 64 trials in a block in random order.

Table 1. All stimulus sounds used in the present study

	Syllable names	Pitches	Syllable name distance	Pitch distance
Congruent	'dol'	C4	Far	Far
		E4	Far	Near
	'mi'	C4	Near	Far
		E4	Near	Near
	'sol'	G4	Near	Near
		B4	Near	Far
	'xi'	G4	Far	Near
		B4	Far	Far
Incongruent	'dol'	G4	Far	Near
		B4	Far	Far
	'mi'	G4	Near	Near
		B4	Near	Far
	'sol'	C4	Near	Far
		E4	Near	Near
	'xi'	C4	Far	Far
		E4	Far	Near

Procedure

Participants were tested on a computer using E-Prime 2.0 software. They sat on a comfortable chair, 70 cm away from a screen, wearing earphones. Two white dots were attached to the table between the participant and the screen; one was about 5 cm away from the participant and the other was about 5 cm away from the screen.

The participants were instructed to hold two symmetric mouses, one in each hand, and to place his/her index fingers on the left or right button. Participants used either a left-hand-front posture (i.e. putting the mouse in his/her left hand on the far dot and the mouse in the right hand on the near dot) or a right-hand-front posture. All participants used both of the postures in the experiment, with the orders of postures counterbalanced between participants.

Participants were informed that they would hear two successive auditory stimuli in each trial, and their task was to decide whether the target sound (the second sound) was higher or lower than the anchor sound (the first sound). At the beginning of each trial, a fixation appeared in the center of the screen for 1000 ms. After the fixation disappeared, the anchor sound appeared in the

https://doi.org/10.1017/prp.2019.25 Published online by Cambridge University Press

Table 2. Summary of response times and accuracy in different pitch and syllable name change conditions

			RT		Accuracy (%)	
	Pitch change	Syllable name change	*M*	*SD*	*M*	*SD*
Congruent	Descending	Descending	1014.45	563.29	90.91	28.76
	Ascending	Ascending	992.47	496.66	92.23	26.77
Incongruent	Descending	Ascending	1082.90	564.92	88.45	31.97
	Ascending	Descending	905.69	486.60	89.58	30.56

earphones for 1000 ms, following a 500 ms silent interval. After the interval, the target sound was presented for 1000 ms. The screen was blank throughout the trial except for during the fixation phase. The trial ended when the participant reacted, followed by the next trial.

Participants responded by clicking one of the mouses in their hands. In half of them, the far mouse indicated that the target was higher than the anchor sound while the near mouse indicated that the target was lower (accordant task mapping). For the other half of them, the mapping was reversed (reversed task mapping). Accuracy and response times were recorded as dependent variables.

There were four 64-trial-long test blocks throughout the experiment, as well as two 12-trial-long training blocks before the first and third test blocks. Participants switched their hand postures before the second training block. The training trials were identical to the test trials except for the feedback immediately shown after they responded. Participants had to reach 80% accuracy in the training block to enter the test block (if not, they had to do the training block again).

Results

We eliminated the trials with incorrect responses (9.71% of the total trials) in the RT analysis. Also, we excluded any RT data that was three standard deviations away (2.2% of the total trials) from the average in trials with correct responses.

Table 2 shows the accuracy and RT in different pitch and syllable name conditions. Participants' mean accuracy in congruent trials was 91.57% ($SD = 27.78\%$, 95% CI [[90.73%, 92.41%]), whereas their mean accuracy in incongruent trials was 89.02% ($SD = 31.27\%$, 95% CI = [88.07%, 89.96%]). Their mean response time in congruent trials was 1003.33 ms ($SD = 530.68$, 95% CI = [986.41, 1020.25]), whereas it was 993.13 ms in incongruent trials ($SD = 534.04$, 95% CI = [975.83, 1010.43]).

The influence of syllable name change over pitch change and the spatial representation of pitches

A 2 × 2 (congruency × task mapping) mixed ANOVA for accuracy data was conducted, with congruency as within-subject variable and task mapping as between-subject variable. This analysis only revealed a significant main effect of congruency, $F(1,31) = 5.22$, $p = .03$, $\eta_p^2 = .14$ [.01–.32], participants were more accurate in congruent trials than in incongruent trials, suggesting the occurrence of interference between the syllable name change and the pitch change. The main effect of task mapping was not significant, $F(1,31) = 0.00$, $p = .99$, $\eta_p^2 = .00$ [.00–.00], nor was the interaction, $F(1,31) = 0.50$, $p = .49$, $\eta_p^2 = .02$ [.00–.14]).

We also ran a similar 2 × 2 (congruency × task mapping) mixed ANOVA for RT. No effect was significant (the main effect of congruency: $F(1, 31) = 0.35$, $p = .56$, $\eta_p^2 = .01$ [.00–.13]; the main effect of task mapping: $F(1, 31) = 1.96$, $p = .17$, $\eta_p^2 = .06$ [.00–.22]; the interaction of congruency and task mapping: $F(1, 31) = 0.09$, $p = .77$, $\eta_p^2 = .00$ [.00–.01]. The absence of the main effect of task mapping in accuracy or RT data indicates that participants did not exhibit the SMARC effect.

The spatial representation of syllable names

To look further into the spatial representational nature of syllable name change (the SMARC-like effect), we ran a 2 × 2 (syllable name change × react-place) repeated ANOVA for accuracy data. The interaction between syllable name change and react-place was not significant, $F(1,32) = .74$, $p = .40$, $\eta_p^2 = .02$ [.00–.16].

We also ran a 2 × 2 (syllable name change × react-place) repeated ANOVA for the RT data. No syllable name change × react-place interaction was found, $F(1,32) = .04$, $p = .85$, $\eta_p^2 = .00$ [.00–.05]. There was only a main effect of syllable name change, $F(1,32) = 43.77$, $p < .001$, $\eta_p^2 = .58$ [.37–.69]. These results showed that participants did not exhibit the SMARC-like effect.

The distance effect of pitches and syllable names

To test the distance effect, two paired sample *t*-tests were conducted using the distance between the pitches of anchors and targets. Results showed that participants were more accurate in the far trials than in the near trials, $t(32) = -10.33$, $p < .001$, Cohen's $d = -1.80$ [−2.35–1.24]; they also responded more rapidly in the far trials than in the near trials, $t(32) = 8.58$, $p < .001$, Cohen's $d = 1.49$ [0.99–1.99]. These results indicate the existence of the distance effect for pitches.

Similarly, two paired sample *t*-tests were conducted using the distance between the syllable names of anchor and target. Results showed no trend in accuracy, $t(32) = -.77$, $p = .45$, or RT, $t(32) = 1.00$, $p = .33$, indicating the absence of distance effect for syllable names.

The sequential effect of syllable names

Because neither the SMARC-like effect nor the distance effect was found in previous analyses for syllable names, we tested the potential sequential effect. A paired sample *t*-test for the RT data was conducted with the sequence positions (early or late in the sequence). Participants' reactions were faster in early syllable name trials than in late syllable name trials, $t(32) = 6.59$, $p < .001$, Cohen's $d = 1.15$ [1.58–0.701].

https://doi.org/10.1017/prp.2019.25 Published online by Cambridge University Press

Discussion

In the present study, we tested our main hypothesis concerning the interference between the syllable name change and the pitch change by examining both accuracy and response time in different syllable name change and pitch change conditions in a pitch comparison task. In line with our main hypothesis, there was a significant congruency effect in terms of accuracy, though not in the RT data. Participants were more accurate when syllable name change and pitch change were congruent compared to when they were not. This suggests that the direction of syllable name change was automatically represented when non-musicians judged the direction of pitch change.

It is interesting that the pitch-syllable name congruency effect was only found in accuracy but not in RT data. This pattern is similar to studies conducted with non-musicians or musically unselected participants. For example, with musically unselected participants, Campbell and Scheepers (2015, experiment 1) similarly observed interference between pitch change and number change in the pitch comparison task only in accuracy but not in RT. However, it is different from the pattern in many auditory Stroop studies. In those studies, AP-possessors and musicians perform almost perfectly in accuracy (Grégoire, Perruchet, & Poulin-Charronnat, 2013; Itoh et al., 2005), but their reaction time to incongruent trials was longer than congruent trials. We speculate that the reason for this pattern difference is two-fold. On the one hand, the close-to-ceiling accuracies of music experts left little space for the analyses of variation. Therefore, the effect mainly appeared in RT data. On the other hand, it is the expertise that results in the difference in behavioural patterns. Being capable of inhibiting the undesired answers, music experts are only affected when generating the right answers. However, non-musicians are more vulnerable to distractions from other dimensions. They are more likely to be led to mistake decisions instead of delayed right answers.

We also examined the possible representations involved in the interference between pitch and syllable name change. The SMARC and SMARC-like effects, the distance effect, and the sequential effect were examined to test if the involved representations were spatial, magnitude, or sequential. First, we examined the association between pitch change and react place (known as the SMARC effect) and the association between syllable name change and react place (the SMARC-like effect). For the SMARC effect, the main effect of task mapping is not significant in the test of accuracy or RT. This result was contrary to our expectations and inconsistent with previous findings (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi et al., 2006). This might suggest that the pitch change did not activate a spatial representation in the pitch comparison task. However, with task mapping as a between-subject variable, we cannot rule out the possibility that the individual difference between the groups was so big that the SMARC effect might be masked in the present study. For the SMARC-like effect, there seems to be no salient spatial representation of syllable name change in the experiment. For tests of both accuracy and RT, we found no significant interaction between syllable name change and react-place. Because syllable name was an irrelevant dimension and balanced in the experiment design, each participant reacted to far and near places for the same times when each syllable name was heard. Therefore, this was a within-subjects design for the interaction of syllable name change and react-place. Even so, no positive result was found. Therefore, we believe that there was no spatial representation involved in the interference of syllable name change and pitch change.

Second, the examinations of distance effects showed a different pattern. There was an evident distance effect for pitch. Participants responded more accurately and more rapidly when the target sound was far from the anchor sound. This result is in line with previous findings (Kadosh et al., 2008). However, there was no distance effect for the change in syllable names for accuracy or response time. These results indicated that the representation of the change in syllable names did not involve a component of magnitude.

Last, for the sequential effect, we found that participants reacted to trials with early syllable names more rapidly than trials with late syllable names. The relationship between RT and syllable name position indicates that a mental searching process exists. As in Halpern's (1988) study, auditory stimuli are arranged in a monotonous sequence. The retrieving of the target unit in the sequence requires scanning through the sequence one-by-one from the very beginning. Therefore, the later the target is in the sequence, the more time it needs to retrieve it. In the present study, participants heard tones sung with syllable names in both anchor and target sounds. Though the anchor sound remained constant, the target sounds varied in their syllable names. Upon hearing the target syllable name, the scanning (from dol to the target syllable name) happened automatically, resulting in the sequence effect in the RT data.

Based on the results mentioned above, we speculate that the representation of syllable name changes is sequential instead of spatial or magnitude. When two syllable names were heard in the experiment, they automatically activated a representation of sequential relation according to their positions in the sequence. Note that the participants were non-musicians who could not identify the fixed relationships between absolute pitches and their corresponding syllable names, which has been shown in many previous studies (Akiva-Kabiri & Henik, 2012; Grégoire et al., 2013, 2014; Schulze et al., 2013; Tsai, Chen, Wen, & Chou, 2015). Therefore, in the experiment, participants did not activate the specific pitch height that each syllable name represented. It was the sequential order of syllable names that was automatically activated and interfered with the judgment of relative pitches. These results indicate that non-musicians can access the relative relationships of two syllable names. More importantly, this process is rather proficient in that it can be carried out spontaneously and automatically among non-musicians.

The present study is the first evidence that syllable name change interferes with non-musicians' judgments of pitch change. Contrary to perceiving absolute pitches or syllable names, non-musicians perceived the relative changes in both pitches and syllable names. When the pitches and syllable names changed in the same direction (congruent trial), participants were more accurate than when they changed in different directions (incongruent trial).

Unlike pitches, syllable names are labels created by culture. They have no natural association with magnitudes from the physical world. However, in the Western music system, syllable names covary with pitches in a given octave, thus forming a monotonously arranged sequence. Through the basic musical theory education in schools, even non-musicians acquire the knowledge of pitch-syllable name covariation and sequential order of syllable names. Their knowledge of syllable names might cause the automatic accessing of sequential orders of syllable names, even when they are not asked to do so. Another issue worth noticing is that cultural influence is so far-reaching that even non-musicians, the majority of the population, represent the sequential order of syllable names automatically. This is another example of cultural products that not only serve but also influence people's perceptual and cognitional activities in daily life.

https://doi.org/10.1017/prp.2019.25 Published online by Cambridge University Press

In conclusion, we examined the existence of interference between the syllable name change and pitch change in a pitch comparison task in non-musicians. Non-musicians' judgments of pitch change were more accurate in congruent trials than in incongruent trials. This result showed that syllable names do influence non-musicians' representation of pitches in a relative manner. We propose that it is the sequential order of syllable name change, which is a product of cultural activities, that activates a sequential representation and interferes with the judgment of pitch change.

Financial support. None.

References

Akiva-Kabiri L. and Henik, A. (2012). A unique asymmetrical Stroop effect in absolute pitch possessors. *Experimental Psychology*, **59**, 272–278.

Campbell J.I.D. and Sacher S.G. (2012). Semantic alignment and number comparison. *Psychological Research*, **76**, 119–128.

Campbell J.I.D. and Scheepers F. (2015). Effects of pitch on auditory number comparisons. *Psychological Research*, **79**, 389–400.

Grégoire L., Perruchet P. and Poulin-Charronnat B. (2013). The musical Stroop effect: Opening a new avenue to research on automatisms. *Experimental Psychology*, **60**, 269–278.

Grégoire L., Perruchet P. and Poulin-Charronnat B. (2014). About the unidirectionality of interference: Insight from the musical Stroop effect. *Quarterly Journal of Experimental Psychology*, **67**, 2071–2089.

Halpern A.R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 434.

Itoh K., Suwazono S., Arao H., Miyazaki K. and Nakada T. (2005). Electrophysiological correlates of absolute pitch and relative pitch. *Cerebral Cortex*, **15**, 760–769.

Kadosh R.C., Brodsky W., Levin M. and Henik A. (2008). Mental representation: What can pitch tell us about the distance effect? *Cortex*, **44**, 470–477.

Kosslyn S.M., Ball T.M. and Reiser B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of experimental psychology: Human Perception and Performance*, **4**, 47.

Lidji P., Kolinsky R., Lochy A. and Morais J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology. Human Perception and Performance*, **33**, 1189–1207.

McDermott J.H. and Oxenham A.J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, **18**, 452–463.

Morgan A.L.R. and Brandt J.F. (1989). An auditory Stroop effect for pitch, loudness, and time. *Brain and Language*, **36**, 592–603.

Patel A.D. (2012). *Music, language, and the brain* (Y. Yang et al., Trans.). Shanghai: Huadong Normal University Press. (Original work published 2008).

Rusconi E., Kwan B., Giordano B.L., Umiltà C. and Butterworth B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, **99**, 113–129.

Schulze K., Mueller K. and Koelsch S. (2013). Auditory Stroop and absolute pitch: An fMRI study. *Human Brain Mapping*, **34**, 1579–1590.

Tsai C.-G., Chen C.-C., Wen Y.-C. and Chou T.-L. (2015). Neuromagnetic brain activities associated with perceptual categorization and sound-content incongruency: A comparison between monosyllabic words and pitch names. *Frontiers in Human Neuroscience*, **9**, 1–10.

https://doi.org/10.1017/prp.2019.25 Published online by Cambridge University Press