Sensorimotor Mismapping in Poor-pitch Singing

Hao He and Wei-dong Zhang, *Shanghai, China*

Summary: Objective. This study proposes that there are two types of sensorimotor mismapping in poor-pitch singing: erroneous mapping and no mapping. We created operational definitions for the two types of mismapping based on the precision of pitch-matching and predicted that in the two types of mismapping, phonation differs in terms of accuracy and the dependence on the articulation consistency between the target and the intended vocal action. The study aimed to test this hypothesis by examining the reliability and criterion-related validity of the operational definitions.
Study Design. A within-subject design was used in this study.
Methods. Thirty-two participants identified as poor-pitch singers were instructed to vocally imitate pure tones and to imitate their own vocal recordings with the same articulation as self-targets and with different articulation from self-targets.
Results. Definitions of the types of mismapping were demonstrated to be reliable with the split-half approach and to have good criterion-related validity with findings that pitch-matching with no mapping was less accurate and more dependent on the articulation consistency between the target and the intended vocal action than pitch-matching with erroneous mapping was. Furthermore, the precision of pitch-matching was positively associated with its accuracy and its dependence on articulation consistency when mismapping was analyzed on a continuum. Additionally, the data indicated that the self-imitation advantage was a function of articulation consistency.
Conclusion. Types of sensorimotor mismapping lead to pitch-matching that differs in accuracy and its dependence on the articulation consistency between the target and the intended vocal action. Additionally, articulation consistency produces the self-advantage.
Key Words: Poor-pitch singing–Pitch-matching–Sensorimotor mapping–Precision–Articulation.

INTRODUCTION

Individuals who are unable to match musical pitch accurately or precisely are considered to have poor-pitch singing.[1,2] Researchers measure the proficiency of pitch singing by using pitch-matching tasks in which participants are instructed to vocally imitate target pitches.[3–6] In terms of objective measurement based on acoustic methods that extract the fundamental frequency ($F0$) of a sound, two measures are used to evaluate the pitch-matching performance. One is accuracy, which refers to the average difference between the target pitch and the pitch that an individual sings. Accuracy scores represent the proximity of vocally produced pitches to the target pitches. Accuracy is measured in cents or semitones.[a] Because the human song system is highly complicated and involves peripheral mechanisms of vocalization and relevant neural networks,[7,8] it is less likely for an individual to exactly match a target pitch. Thus, an acceptable range of pitch deviation must be determined for accurate pitch-matching. The cutoff for acceptable deviation of a produced pitch from the target pitch ranges from 100 cents[2,5] to 50 cents.[3,4,6] Another measure for pitch-matching is precision. Precision, which in statistics refers to the standard deviation (SD) of sung pitches, measures the variability of pitch-matching across repeated trials.

The cutoff that defines an acceptable range of precision scores for precise pitch-matching is set at 100 cents or 50 cents.[2,3]

Mechanism of poor-pitch singing: background

According to previous models of the human song system,[7,9–11] poor-pitch singing can be caused by a dysfunction of perception,[12–18] sensorimotor integration,[4–6,19] vocal motor control,[4,20,21] or memory.[15,22] Among these, a malfunction in sensorimotor integration is thought to be the main cause of poor-pitch singing.[4,5]

Regarding the mechanism of sensorimotor dysfunction in poor-pitch singing, an early explanation suggested that sensorimotor mismapping between pitch percepts and phonatory gestures caused systematic singing errors. The explanation accounted for the phenomenon of interval compression that occurred while singing melodies,[5,23] but it was challenged based on the considerable variability in pitch-matching.[2–5] Given this limitation, researchers introduced the internal model framework based on the domains of motor planning and control,[24] and proposed the inverse modeling mechanism, arguing that poor-pitch singers lack sensorimotor translation between pitch percepts and phonatory gestures.[6] However, the inverse modeling account was not more informative than the previous sensorimotor mismapping account in terms of explaining how the sensorimotor mismapping forms.

A recent multimodal imagery association (MMIA) model offered some answers to this question.[19] In the MMIA model, comfort pitch, the pitch that an individual sings comfortably, plays a role in distorting the initially unbiased sensorimotor mapping via its attracting influence on pitch-matching. Mathematically, the distortion is explained as the convolution of two probability distributions: (1) an unbiased sung pitch distribution that centers on the target pitch and (2) a comfort pitch distribution. The MMIA model explains not only systematic bias but also the

large variability in pitch-matching that was observed in previous studies.[3–5]

The MMIA model, in a mathematical way, sheds light on the formation mechanism of sensorimotor mismapping in poor-pitch singing. However, little research has elaborated the construct of the sensorimotor mapping system. In fact, the four categories of pitch-singing derived from the combination of accuracy (accurate *vs.* inaccurate) and precision (precise *vs.* imprecise), which have been systematically examined in previous studies,[2,3] suggest that there may be multiple types of sensorimotor mapping. In particular, because pitch-matching can be either imprecise (both inaccurate and accurate) or inaccurate yet precise, there are likely different types of mismapping.

A new hypothesis

With the aim of exploring the construct of mismapping at the category level, this study proposes the hypothesis that sensorimotor translation involves a mapping system that encompasses three types of mapping: accurate mapping, erroneous mapping, and no mapping. Accurate mapping typically leads to accurate and precise pitch-matching. With erroneous mapping, an individual forms a fixed but inaccurate connection between a pitch percept and a phonatory gesture that does not produce the pitch. When erroneous mapping occurs, the sung pitch varies little from a certain pitch height (but is not the same as the target), producing inaccurate but precise pitch-matching. With no mapping, there is no formed connection between a pitch percept and a phonatory gesture. In such cases, an individual may randomly map a pitch percept onto a phonatory gesture each time he or she imitates the target, producing imprecise pitch-matching across repeated trials.

This study then proposes an operational definition of each type of mapping based on the relationship between accuracy and precision. Accuracy is measured as the absolute value of the mean signed pitch deviation, whereas precision is measured as the SD of signed pitch deviation. Because a specific cutoff value may be more sensitive to precision than accuracy,[2,3] we established a 50-cent cutoff for accuracy and a more liberal cutoff of 100 cents for precision in this study. Using these cutoff values, accurate and precise pitch-matching represents correct auditory-motor mapping; inaccurate but precise pitch-matching represents erroneous mapping; and imprecise pitch-matching, regardless of whether it is accurate, represents no mapping. It is worth noting that an individual can exhibit both types of sensorimotor mismapping simultaneously or just one of them.

Hypothesis testing

The proposed sensorimotor mismapping hypothesis can be tested by examining the reliability and validity of the operational definitions of the types of mismapping. For reliability testing, this study used the split-half approach, which tests the degree of correlation between data of the first and the second halves.[25] In particular, we analyzed reliability by splitting trials into two equal parts in terms of time order and investigating the correlation between the parts.

To assess validity, we used criterion-related validity, which assumes that there is a correlation between a test measure and one or more external criteria.[26] Two criteria were applied in this study. The first criterion was the accuracy of pitch-matching. Pitch-matching with erroneous mapping should be more accurate than pitch-matching with no mapping for two possible reasons. First, erroneous mapping may be the outcome of a compromise after multiple failed attempts to match a target pitch. Although individuals may know that they make singing errors, they may be unable at the time to improve their singing accuracy; thus, they must accept their most accurate vocal action for the sake of expediency. Considering that the majority of poor-pitch singers have normal pitch perception,[4,5] the produced pitch deviation would not be considerable in such cases. Second, erroneous mapping can result from limited perceptual resolution of human vocalizations. Previous research has shown that the differential threshold is higher when the targets are natural human voices than when the targets are synthesized vocal sounds.[4] Thus, for individuals who have normal pitch perception, the failure to detect deviations in the produced pitch through auditory feedback can lead to erroneous mapping, but their accuracy would not be substantially affected. In contrast, no mapping can cause large amplitudes of pitch deviations, as in cases when individuals attempt to match pitches by trial and error and still do not succeed.

The second criterion for the validity test was the dependence of phonation on the articulation consistency between the target and the intended vocal action. Our starting point was a widely recognized finding that phonation and articulation are structurally linked. Early research has shown that a change in articulation changes the produced pitch.[27–29] Additionally, the articulation consistency between the target and the vocal action can affect the accuracy of pitch-matching. As research has shown, pitch-matching is more accurate when human voices are imitated than when tones produced by instruments are imitated,[4,30–32] and it is the most accurate in self-imitation.[4,6,31] This is the so-called human voice or self-imitation advantage. Note also that the timbre similarity explanation has been ruled out by previous research.[6] The increased accuracy of pitch-matching when imitating human voices and during self-imitation suggests that phonation is affected by the articulation consistency between the target and the intended vocal action. In other words, the human voice or self-imitation advantage may be a function of articulation consistency.

In the sensorimotor mismapping hypothesis proposed in this study, phonation is more dependent on the articulation consistency in no mapping than in erroneous mapping for the following reasons. In pitch imitation, a phonatory solution is needed when an auditory input is received. In erroneous mapping, a phonatory motor plan can be quickly generated because of the sensorimotor mappings that are formed. In no mapping, the planning of phonation is not directly driven by a formed sensorimotor association, but it can be affected by external vocal motor cues. Articulation, a set of complicated and rapid motor behaviors involving the coordination of laryngeal, pharyngeal, and orofacial muscles, can provide such cues. Articulation is more imitable than phonation because articulation, which comprises such components as lip, tongue, and jaw movements, can be observed, whereas phonation lies deep in the throat and is difficult to observe. Thus, the unplanned phonation in no mapping is more