

Linguistic limitations of describing sound: Is talking about music like dancing about architecture?

Helen F. Mitchell¹ and Raymond A. R. MacDonald²

¹ Sydney Conservatorium of Music, University of Sydney, Australia

² Department of Psychology, Glasgow Caledonian University, UK

Verbalizing sound quality presents a challenge to musicians and pedagogues in describing a complex sensory phenomenon. Verbal descriptions may only be effective when a performer's sound translates easily and completely into words. A verbal overshadowing (VO) effect may occur when a verbal description distorts the recall of the original aural memory. The aim of this study is to examine the impact of verbal overshadowing in a music performance context. This exploratory pilot project builds on VO research in other fields to assess the value and limitations of language in describing performers' sound quality. Outcomes will have implications for musicians and pedagogues in the use of language in music teaching and learning.

Keywords: music perception; verbalization; auditory recognition; music performance; singing voice

Verbal overshadowing (VO) can occur when we use words to describe sensory experiences (such as sight, taste, or hearing). Numerous studies have examined this phenomenon in face, taste, and sound recognition and report that verbal description, or verbal encoding, impairs later recall of the respective experience. Individuals were less likely to identify the target stimuli from a line-up (Parr *et al.* 2002, Perfect *et al.* 2002, Schooler and Engstler-Schooler 1990). The perception and description of music performance presents a similar sensory challenge. Can words adequately capture sound quality in a music performance? And if so, which terms or group of terms are most effective in expressing sound quality?

The VO effect has been most apparent in face recognition, where witnesses to crimes have to describe the perpetrator of a robbery in detail and

later select their face from a photographic line-up. The act of verbal description has negatively influenced individuals' ability to recognize the face they described in words (Schooler and Engstler-Schooler 1990). Empirical studies have confirmed this VO effect in facial recognition by comparing verbal description groups with memory-only control groups. More recent studies have demonstrated that there is a similar VO effect on listeners' ability to recognize spoken voices following verbal description while no-description control groups were able to isolate the voice from a line-up (Perfect *et al.* 2002, Vanags *et al.* 2005). Interestingly, expert wine-tasters were more likely than novices to recognize wine-relevant odors by their verbal labels (Parr *et al.* 2002) despite the fact all tasters had knowledge of language used to describe wine. Novices may have lacked the ability to separate perceptual and linguistic skills while experts' superior perceptual skills protected them from linguistic interference to taste recognition.

In music, the modality of the stimuli (aural) does not match the task (verbal description), yet we rely on verbal descriptors to explain and communicate our perception of sound quality in training and assessment. There is growing interest in the way listeners perceive, rate, and describe music performances. Like the witnesses above, listeners form a global impression of the performance or performer and use a limited selection of verbal descriptors to explain the reasons for their judgments. They focus on the more easily verbalized technical and performance components rather than describing overall quality (Davidson and Coimbra 2001, Stanley *et al.* 2002). In fact, examiners have even noted candidates' dress and stage manner to facilitate later recall of individual singers' performances for discussion and assessment (Davidson and Coimbra 2001). It is more natural to conceptualize sound as belonging to a sound-producing object, rather than analyze the components that make up the sound (Ekholm *et al.* 1998). The use of language may effectively hinder our recall of performers and performances.

While there is a growing body of knowledge focused on categorizing terminology used in describing vocal quality, there is still much to be learned about verbal descriptors in listeners' perception of music performance. Communication may focus on the most easily verbalized characteristics in the sound rather than individual composition of these characteristics that makes a sound unique (Kenny and Mitchell 2006, Mitchell and Kenny 2006). Verbalization, rather than enhancing interpretation of sound qualities, may have a disruptive effect on listeners' recall and communication of music performance. The main objective of this project is to examine impact of verbal description on the recall of music performances and, specifically, singers. It investigates whether verbal overshadowing (VO), or putting a sound "into

words,” distorts listeners’ memory and subsequent recall of the original performance or performer. This paper presents work in progress, the full details of which will be reported in subsequent publications.

METHOD

Participants

Singers and listeners were recruited from the staff and students at the Sydney Conservatorium of Music (SCM).

Materials

Singers (n=6) performed two short song excerpts in a sound-treated studio at SCM and recorded using a matched pair of stereo microphones (Neumann KU100, ORTF configuration). Samples of their recordings were played to listeners in the perceptual test.

Procedure

Listeners (n=24) were asked to attend a single listening session at SCM and were informed that the purpose of the study was to investigate the effect of visual and verbal tasks on the ability to recognize aural stimuli. Audio samples were played to listeners from a CD player via circum-aural closed-back stereo monitoring headphones (Sennheiser HD 270) to ensure that each listener heard the same quality of sample. Listener participants were assigned to either a *verbal description* group or a *non-description* control group.

Listeners were presented with a short sample of a single singer performing *song 1* (for encoding) on CD. Singer presentation was randomized from the six singers. Listeners were then asked to perform a visual maze task on paper (a filler-task) for 10 mins. In the following 5 mins, the experimental group was asked to write a detailed description of the voice they heard at the start of the test, while the control group completed an anagram puzzle. Finally, both groups were presented a line-up of six voices singing a short excerpt from *song 2* and were asked to identify the original voice they heard and to rate their confidence in their response on a scale of 1-10 (1=not confident, 10=very confident).

RESULTS

Listeners’ responses were examined for correct voice identification by each group (verbal description and non-description control). Ten of 24 listeners

Table 1. Correct and incorrect voice line-up identifications, and mean confidence scores for the verbal description and non-description groups.

	<i>Verbal description group</i>	<i>Non-description control group</i>
Correct identification	4	6
Incorrect identification	8	6
Mean score (SD) for confidence in choice	5.7 (2.1)	6.3 (2.2)

correctly identified their target voice from the line-up. Table 1 shows the distribution of responses by experimental group. Fifty percent of the control group and 33% of the verbal description group correctly identified their target voice.

Verbal descriptions

For the written descriptions of target voices, listeners were instructed to focus on the characteristics of the voice singing the target song rather than the song the singer sang. Most listeners identified that the singer was a soprano, although some listeners debated whether their target singer might be a mezzo soprano, with comments such as: “slightly darker sound than for soprano” and “voice in a lower timbre to that of a soprano.” Eight listeners mentioned their target singer’s use (or lack) of vibrato—for instance, “singer does not use an abundance of vibrato,” “smooth even vibrato on long notes but none on shorter notes,” and “vibrato enlarged at the end of phrase; vibrato change may be for emotional purpose?”.

Seven listeners mentioned the singer’s diction or articulation of language: “good diction,” “the articulation of the words was extremely clear,” and “when singing, the words are clearly understandable.” Listeners made general judgments about the quality of their target voices without expressing the reasons that motivated their statements: “not very expressive at all, after thinking about it; quite boring,” “not the style of music I listen to but singer is good,” “quality of the instrument has potential, needs refinement of technique to bring out the individual voice, but she is still young,” and “sings in an unemotional way (possibly she doesn’t know/understand what she’s saying).” Some listeners mentioned words usually associated with classical singing voice, such as “rich and smooth,” “bell-like,” “round,” and “covered” while others used technical descriptors like “support” and “projection.” For two listeners, the sound they heard evoked a mental picture of the target voice

singing: “English accent/Anglo-Saxon, blonde, quite tall” and “image evoked based on sound: larger brunette female...solid, fairly tall.”

DISCUSSION

This is the first study to investigate the effect of VO in the perception of music performers. Verbal description decreased listeners’ likelihood of identifying the target voice but only showed a slight reduction in listeners’ confidence in their identification choice. Listeners were less likely to pick the correct voice but were not aware that verbalizing their perception of the target voice reduced their ability to achieve correct identification.

Listeners’ responses in this study follow a similar test of spoken voice recognition (Perfect *et al.* 2002), where only 50% of listeners in the non-verbal group and 21.4% of the verbal description group correctly identified the target voice from a line-up. This project will confirm if music listeners are similarly susceptible to the effects of VO when they describe singing voices.

Listeners’ descriptions of the target voices illustrated the limits of language to communicate music perception. Examination of the comments further confirmed that vocal quality is difficult to articulate or itemize (Davidson and Coimbra 2001, Ekholm *et al.* 1998). While verbalizing sound quality in this context proved an unusual challenge, listeners still resorted to descriptions of the song, or indeed of the person singing, rather than attempt to describe the idiosyncrasies of each voice.

These preliminary results form part of a more extensive study investigating the nature of VO effects in music listeners and the extent to which the results from VO studies in other domains generalize to the musical context. Moreover, this ongoing work is also highlighting the complexities and ambiguities in describing singing voices linguistically.

Acknowledgments

The first author was funded by an early career researcher grant from the Sydney Conservatorium of Music, University of Sydney. We wish to thank the singers and listeners who took part in this study.

Address for correspondence

Helen Mitchell, Sydney Conservatorium of Music, Building C41, University of Sydney, Sydney, New South Wales 2006, Australia; *Email:* h.mitchell@usyd.edu.au

References

- Davidson J. W. and Coimbra D. (2001). Investigating performance evaluation by assessors of singers in a music college setting. *Musicae Scientiae*, 5, pp. 33-53.
- Ekholm E., Papagiannis G. C., and Chagnon F. P. (1998). Relating objective measurements to expert evaluation of voice quality in Western classical singing: Critical perceptual parameters. *Journal of Voice*, 12, pp. 182-196.
- Kenny D. T. and Mitchell H. F. (2006). Acoustic and perceptual appraisal of vocal gestures in the female classical voice. *Journal of Voice*, 20, pp. 55-70.
- Mitchell H. F. and Kenny D. T. (2006). Can experts identify “open throat” technique as a perceptual phenomenon? *Musicae Scientiae*, 10, pp. 33-58.
- Parr W., Heatherbell D., and White K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, 27, pp. 747-755.
- Perfect T. J., Hunt L. J., and Harris C. M. (2002). Verbal overshadowing in voice recognition. *Applied Cognitive Psychology*, 16, pp. 973-980.
- Schooler J. W. and Engstler-Schooler T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, pp. 36-71.
- Stanley M., Brooker R., and Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education*, 18, pp. 43-52.
- Vanags T., Carroll M., and Perfect T. J. (2005). Verbal overshadowing: A sound theory in voice recognition? *Applied Cognitive Psychology*, 19, pp. 1127-1144.