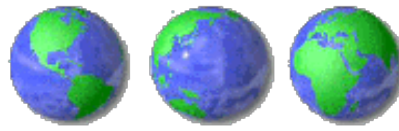


156th ASA Meeting

Miami, FL



[[Lay Language Paper Index](#) | [Press Room](#)]

The Speech-to-Song Illusion

Diana Deutsch

Department of Psychology
University of California, San Diego
La Jolla, CA 92093
ddeutsch@ucsd.edu

Rachael Lapidis

Department of Psychology
University of California, San Diego
La Jolla, CA 92093
rlapidis@ucsd.edu

Trevor Henthorn

Department of Psychology
University of California, San Diego
La Jolla, CA 92093
trevor@music.ucsd.edu

Lay-language version of paper [2aMU6](#)

Presented at 11.15 a.m. on Tuesday, November 11, 2008 in PLUM A & B
156th Meeting of the Acoustical Society of America

This paper reports the first formal investigation of a striking illusion: A spoken phrase is made to be heard convincingly as sung rather than spoken, and this perceptual transformation occurs without altering the signal in any way, or adding any musical context, but simply by repeating the phrase several times over. The illusion is surprising, as it is generally assumed that whether we perceive a phrase as spoken or as sung depends on the physical characteristics of the sound.

The phrase occurs in a sentence in the opening commentary of the compact disc *Musical Illusions and*

Paradoxes (Diana Deutsch, 1995). When you listen to this sentence in the normal way, it appears clearly to be spoken – as indeed it is. Yet when you play a phrase that’s embedded in it several times over, a curious effect emerges: At some point, instead of appearing to be spoken, the phrase appears to be sung.

Here is the full sentence followed by the phrase played repeatedly.

[LISTEN: Sound Demo 1.](#)

And here is the phrase as it is generally heard after it has been played repeatedly:

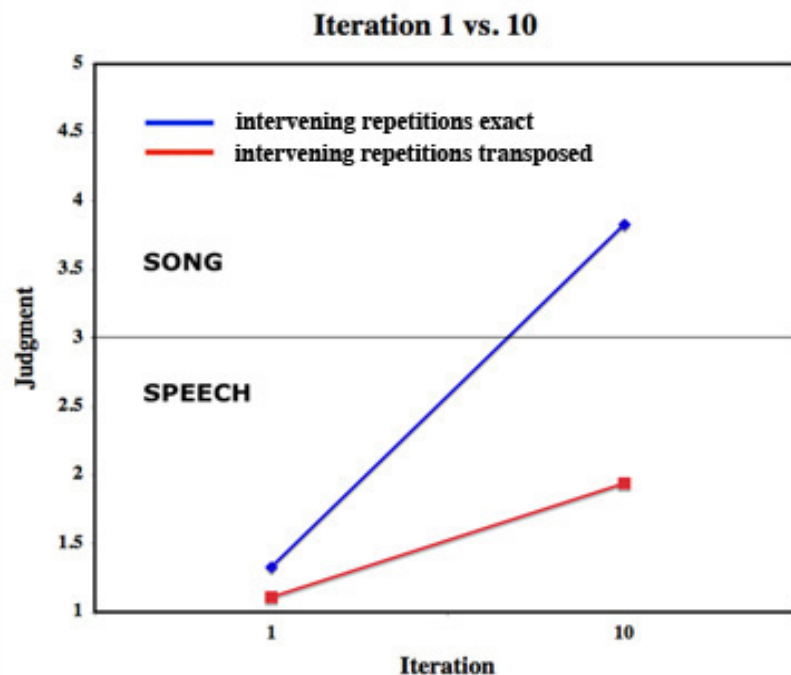


Now here again is the exact same sentence as you just heard . You will probably find that it begins by sounding as speech, just as before. But when you come to the phrase that had been repeated, it suddenly appears to burst into song.

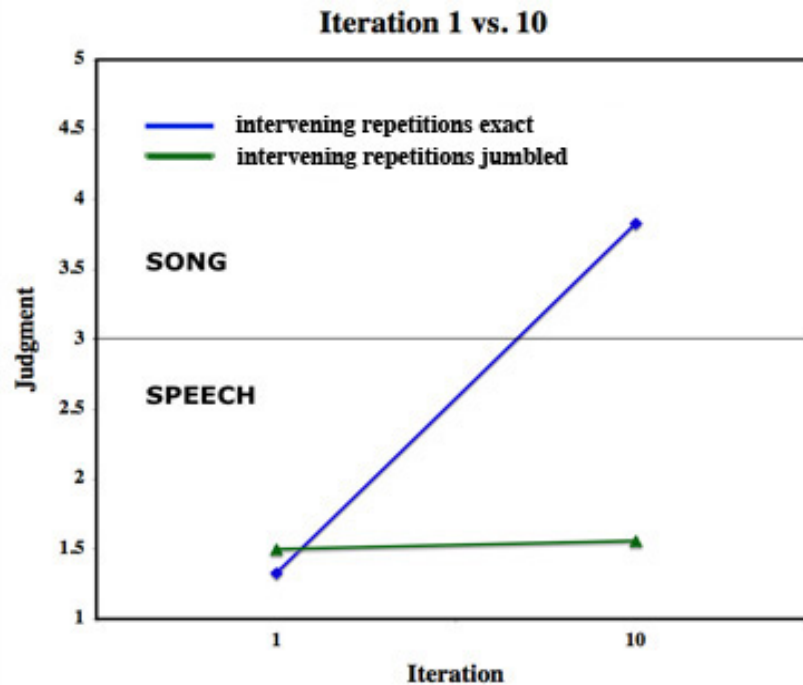
[LISTEN: Sound Demo 2.](#)

In our first experiment we tested three matched groups of subjects, and presented each group with a different condition. The subjects all listened to the full sentence and then to ten presentations of the phrase. During each pause between presentations they judged on a five-point scale whether they heard the phrase as exactly like speech, like speech, like either speech or song, like song, or exactly like song.

In all conditions, the first and last presentations were identical, and we examined the effects of two manipulations of the intervening presentations on the subjects’ judgments. In the first condition, the intervening presentations were exactly as the original. In the second, they were transposed slightly, so that the pitches differed but the pitch relationships were preserved. In the third, the intervening presentations were not transposed, but the syllables were presented in jumbled orderings.



The above graph compares the effects of having the intervening repetitions exactly as the original, as compared with being transposed slightly. As can be seen, when the repetitions were exact, perception moved solidly from speech to song. However, when the repetitions were transposed slightly, although ratings moved slightly towards song, they remained solidly in the speech region.



The above graph shows the effect of having the intervening repetitions consist of the same syllables in jumbled orderings, again compared with having the repetitions exactly as the original. We can see that here there was no transformation from speech to song. So it seems that, in order for this transformation to occur, the phrase needs to be repeated exactly, without transposition, and without changing the ordering of the syllables.

So we can then ask: What do the subjects actually hear when they say that they are hearing song? To find out, we recruited 11 female subjects who had had experience with singing in choirs or choruses, and tested each subject in isolation from the others. We had them listen to the full sentence and then to the phrase repeated ten times, and asked them to reproduce the phrase exactly as they had heard it.

Here are the reproductions of six of the subjects played in sequence. As is evident, although the phrase was spoken, the subjects reproduced it as song.

[LISTEN: Sound Demo 3.](#)

And here are the reproductions of all 11 subjects, digitally mixed together so that they are played as a chorus. (A small amount of reverberation has been added, but otherwise the sounds are exactly as they were recorded.)

[LISTEN: Sound Demo 4.](#)

But one might then wonder whether these subjects could have heard the phrase as sung the first time they heard it. So we recruited another set of 11 subjects on the same basis, and also tested them in isolation from each other. This time we played them the full sentence followed by the phrase presented only once, and asked them to reproduce the phrase exactly as they heard it. Here are the reproductions of six of these subjects played in sequence:

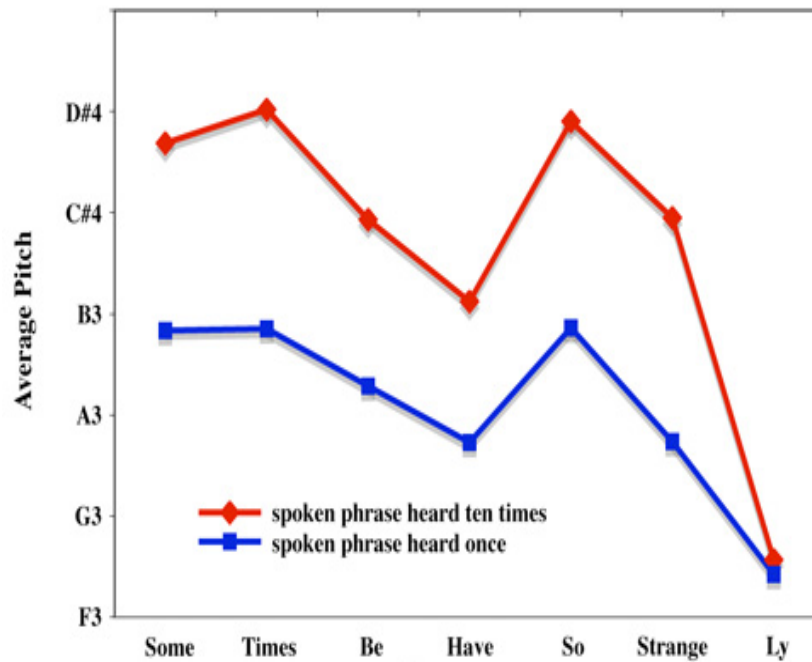
[LISTEN: Sound Demo 5.](#)

And here are the reproductions of all 11 subjects, again digitally mixed together so that they are played as a chorus. This confirms our finding from the rating experiment that when the phrase is heard only once, it is perceived as speech rather than song.

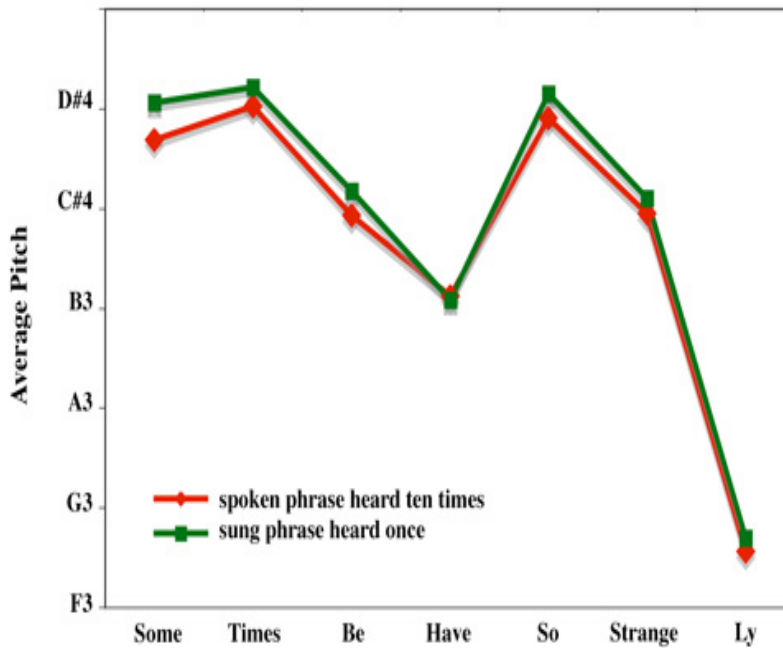
[LISTEN: Sound Demo 6.](#)

To make sure that these subjects were able repeat the pitches after a single hearing, we then had them listen only once to the phrase as sung rather than spoken, and again asked them to repeat back exactly what they had heard. Here are the reproductions of the same six subjects that you just heard, and you can see that they had no problem reproducing the sung melody:

[LISTEN: Sound Demo 7.](#)



The red line in the above graph shows the average pitch of each syllable, averaged over the 11 subjects who repeated back the spoken phrase after having heard it 10 times. The blue line shows the average pitch of each syllable, averaged over the other set of 11 subjects, who repeated back the same spoken phrase after having heard it only once. As can be seen, the reproductions of the two groups were very different.



The red line in the abovegraph again shows the average pitch of each syllable in the *spoken* phrase, averaged over the 11 subjects who repeated it back after having heard it 10 times. The green line shows the average pitch of each syllable, averaged over the other set of 11 subjects, who repeated back the *sung* phrase when it had been presented only once. Notice that there is a remarkable correspondence between these two plots, showing that the subjects' perceptions of the sung phrase were very similar to those of the subjects who had instead heard the spoken phrase repeated 10 times, and quite different from their own perceptions of the spoken phrase when they had heard it only once.

To conclude, this illusion is in line with what philosophers and musicians have been arguing for centuries, that strong linkages must exist between speech and music. We still need to determine the neural processes that are responsible for this striking perceptual transformation. However, the present experiments show that for a phrase to be heard as spoken or as sung, it does not need to have a set of physical properties that are unique to speech, or a different set of physical properties that are unique to song. Rather, we must conclude that, assuming the neural circuitries underlying speech and song are at some point distinct and separate, they can accept the same input, but process the information in different ways so as to produce different outputs. As a further point, this illusion demonstrates a striking example of very rapid and highly specific perceptual reorganization, so showing an extreme form of short - term neural plasticity in the auditory system.