

# DOES REPEATED EXPOSURE TO A TARGET VOICE REDUCE THE IMPACT OF A SIMILAR VOICE?

Pär Wretling<sup>†</sup>, Kirk P. H. Sullivan<sup>†</sup> and Frank Schlichting<sup>‡1</sup>  
<sup>†</sup>Umeå University, Sweden, <sup>‡</sup>University of Hamburg, Germany

## ABSTRACT

It has been demonstrated that a high-quality imitation of a voice can lead to confusion when part of a voice line-up and that discrimination ability does not increase as a consequence of repeated exposure to an imitation of a familiar voice. In this paper the imitation is used only as a distracter in the line-up. The question examined is whether there is a reduction in the impact of the imitation due to repeated exposure. When the imitated voice was the target voice no significant change in response was detected. However, when the imitator's natural voice was the target voice a significant difference in the impact of the similar voice was detected. The degree of variation in the first block responses can be seen as reflecting the security of voice selection. The weaker this is, the greater the possibility for change due to repeated exposure to a target voice.

## 1. INTRODUCTION

The voice line-up is a forensic phonetics technique which is the aural analogue of the police identification parade. In extreme cases as pointed out in [1], the only tangible piece of evidence available to the court might be that of an earwitness. In these cases accurate aural-perceptual voice identification (cf. misidentification) by the earwitness is crucial not only to the success of the prosecution case, but also to the defense of the innocent accused. The earwitness must not positively identify an incorrect voice from the set of different speakers making up the voice line-up.

Schlichting and Sullivan [2] investigated whether imitation can pose a problem for speaker discrimination within the voice line-up. The voice chosen for their experiment was that of a well-known Swedish politician, Carl Bildt (CB). A professional imitator provided an imitation of the voice samples presented in the experiment. Sullivan and Schlichting conducted four different recognition tasks using exactly the same set of sixteen line-ups presented in precisely the same randomized order. The variable was the target voice (the voice to be identified). In the first experiment the target voice was the real voice. In the second experiment, the real voice from long-term memory (the listeners were not presented with a voice, but asked to identify the real voice from memory). In the third experiment the target voice was the imitation (IB) and in the fourth the natural voice of the imitator (IN). The results of Schlichting and Sullivan's experiments indicated that listeners are able to discriminate between the real voice, i.e. the voice being imitation, and the imitation when both are present in a voice line-up. However, the imitation lead to 100 per cent speaker misidentification in the worse case.

Schlichting and Sullivan [3] extended their earlier study by investigating whether repeated exposure to the imitation of Carl Bildt as the target voice (Schlichting and Sullivan's Experiment

Three) had an effect on speaker discrimination within the voice line-up. This study failed to show that the number of correct voice selections changed significantly due to increased exposure to the target voice. Further the listeners failed to recognize the connection between the imitation (IB), and the imitator's natural voice (IN). Thus, even after repeated exposure to the imitation the imitator had managed to keep the person imitated (CB) as the target in the minds of the listeners.

This paper extends Sullivan and Schlichting's work by examining whether repeated exposure to a target voice can reduce the distracting influence of the imitation voice upon the listeners' selection of a voice from the voice line-up containing an imitated voice. Unlike in Sullivan and Schlichting [3], the imitation here is used only as a distracter, that is as a member of the set of voices used in the voice line-ups, and not as the target voice.

## 2. PROCEDURE

The experimental procedure used in this study is identical to that used by [2]. This experiment differs from [2] in that the listeners ability to discriminate and identify the target voice was tested three times. Each test occurred after exposure to the target voice. The experiment was conducted at the same time and using the same listeners as the experiment reported in [2].

### 2.1. Experimental participants

The listeners were randomly selected native Swedish speakers between the ages of seventeen and fifty-seven with no known hearing damage. For each experiment there were 10 female and 5 male listeners. They all indicated that they were familiar with Carl Bildt's voice and not familiar with any of the other speakers.

### 2.2. Speech material

The voice line-ups in this study were constructed from a set of ten voices. These were the voice of Carl Bildt, the former Swedish *statsminister* (Prime Minister), a professional imitation of the voice of Carl Bildt by Göran Gabrielsson, the natural voice of Göran Gabrielsson, three amateur imitations of Carl Bildt, the natural voices of the three imitators and one extra voice. The stimuli presented were *och därför tycker jag att det är så underligt*, 'and therefore I find it so strange' and *att där vill miljöpartiet bromsa*, 'that, there the Green Party wants to drag its heels'. The basis for the selection of these stimuli is presented in [2]. The speech material used to familiarize the listeners was a 30 second extract from a speech Carl Bildt gave as *statsminister* to the *Riksdag*, the Swedish parliament. A transcript is given in the Appendix to [2].

The line-ups comprised six voices. Each voice was separated by a less-than-one-second pause: the pause was around 80 msec in length. Line-ups were constructed with and without

Carl Bildt’s voice, with and without the professional imitation of Carl Bildt’s voice, and with and without the natural voice of the professional imitator. Stimuli, randomly selected from the set clipped from the amateur imitations and their natural voice recordings, were used to make all the line-ups six voices in length. Eight different line-up compositions were, thus, created for each of the two stimuli. This made a total of sixteen different line-ups. The order of presentation of the within the line-up was random. The sixteen different line-ups were also presented in random order.

### 2.3. The recognition tasks

Three different recognition tasks were conducted; these were identical Experiments One, Two and Four presented in [2] The participants heard a one minute recording of the voice they were to identify once before a training block of four line-ups and once before each of the three experimental blocks of sixteen different line-ups. The participants were told that they would hear a voice which they were to remember and identify in the voice line-ups which followed. They were to register, after hearing all six voices, their choice by circling the number indicating the position of the voice in the line-up sequence and if the voice was not

present to mark the not-present option. There was a brief break between each of the experimental blocks.

In Experiment One the target voice was voice of Carl Bildt (CB), and in Experiment Four it was the natural voice of the professional imitator (NI). The task in Experiment Two was slightly different in that the listeners were asked to recognize the voice of Carl Bildt (CB) from memory. This task was included in Schlichting and Sullivan’s original study to confirm that the Swedish population were familiar with Carl Bildt’s voice from memory. Here this task was included as a control; that is to assess to what extent repeating the task itself, without repeated exposure to the target voice, leads to a change in listener response.

### 3. RESULTS

The results of the three experiments are presented below in Tables 1–9 in the form of confusion matrices. Each matrix shows the responses for line-ups 1–16; “—” indicates that a particular voice was not present in a line-up. For each line-up there were 15 listeners. CB = the voice of Carl Bildt (the imitated voice), IB = the imitation of Carl Bildt’s voice; IN = the natural voice of the imitator; F = Foil and NP = Not Present. The correct response is shaded in the matrices.

Line-up	Identified as:					Line-up	Identified as:				
	CB	IB	IN	F	NP		CB	IB	IN	F	NP
1	14	1				9	—	5		1	9
2	13	2				10	—	11			4
3	13			2		11	—			1	14
4	13	2				12	—	13			2
5	14	—			1	13	—	—		1	14
6	14	—		1		14	—	—		3	12
7	14	—			1	15	—	—			15
8	15	—				16	—	—		2	13

Table 1. The confusion matrix for line-ups 1–16. Experiment 1; First experimental Block

Line-up	Identified as:					Line-up	Identified as:				
	CB	IB	IN	F	NP		CB	IB	IN	F	NP
1	14				1	9	—	3			12
2	13	2				10	—	10			5
3	13			1	1	11	—	2			13
4	13	2				12	—	11		1	3
5	13	—			2	13	—	—		1	14
6	15	—				14	—	—		2	13
7	14	—			1	15	—	—			15
8	15	—				16	—	—		2	13

Table 2. The confusion matrix for line-ups ups 1–16. Experiment 1; Second experimental Block

Line-up	Identified as:					Line-up	Identified as:				
	CB	IB	IN	F	NP		CB	IB	IN	F	NP
1	15					9	—			1	14
2	13	2				10	—	9			6
3	14				1	11	—	1		1	13
4	15					12	—	9			6
5	14	—			1	13	—	—			15
6	15	—				14	—	—		2	13
7	14	—			1	15	—	—		1	14
8	15	—				16	—	—		1	14

Table 3. Confusion matrix for line-ups 1–16. Experiment 1; Third experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	14	1				9	—	9		2	4
2	12	3				10	—	13		1	1
3	15		—			11	—	9	—	1	5
4	8	6	—	1		12	—	13	—		2
5	15	—				13	—	—		2	13
6	13	—	1	1		14	—	—	2	8	5
7	14	—	—	1		15	—	—	—	2	13
8	15	—	—			16	—	—	—	1	14

Table 4. Confusion matrix for line-ups 1–16. Experiment 2; First experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	14	1				9	—	9		1	5
2	7	7			1	10	—	14		1	
3	13		—	1	1	11	—	7	—	2	6
4	13	2	—			12	—	13	—		2
5	15	—				13	—	—			15
6	15	—				14	—	—		3	12
7	14	—	—	1		15	—	—	—	1	14
8	15	—	—			16	—	—	—	5	10

Table 5. Confusion matrix for line-ups 1–16. Experiment 2; Second experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	10	3		1	1	9	—	8			7
2	9	6				10	—	10		1	4
3	12	3	—			11	—	7	—	1	7
4	14	1	—			12	—	12	—		3
5	15	—				13	—	—		2	13
6	15	—				14	—	—		1	14
7	15	—	—			15	—	—	—	1	14
8	15	—	—			16	—	—	—	2	13

Table 6. Confusion matrix for line-ups 1–16. Experiment 2; Third experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	3	1	10	1		9	—	4	8	2	1
2	3	1	7	1	3	10	—	3	8		4
3	7	2	—	1	5	11	—	9	—	1	5
4	4	1	—	1	9	12	—	3	—		12
5	4	—	7		4	13	—	—	10	1	4
6	3	—	8	1	3	14	—	—	7	1	7
7	6	—	—	2	7	15	—	—	—	1	14
8	3	—	—	1	11	16	—	—	—	1	14

Table 7. Confusion matrix for line-ups 1–16. Experiment 4; First experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	3	1	10		1	9	—	3	11		1
2	3		11		1	10	—	3	10		2
3	4	2	—	1	8	11	—	6	—		9
4	3	1	—		11	12	—	1	—		14
5	3	—	10		2	13	—	—	12		3
6	3	—	11		1	14	—	—	12		3
7	3	—	—		12	15	—	—	—		15
8	3	—	—		12	16	—	—	—		15

Table 8. Confusion matrix for line-ups 1–16. Experiment 4; Second experimental Block

Line-up	CB	Identified as:				Line-up	CB	Identified as:			
		IB	IN	F	NP			IB	IN	F	NP
1	2		12		1	9	—		11		4
2	3		11		1	10	—	2	11		2
3	3	1	—		11	11	—	7	—	1	7
4	3	2	—		10	12	—	1	—		14
5	3	—	11		1	13	—	—	10	1	4
6	3	—	11		1	14	—	—	11		4
7	4	—	—		11	15	—	—	—		15
8	3	—	—		12	16	—	—	—		15

Table 9. Confusion matrix for line-ups 1–16. Experiment 4; Third experimental Block

### 3.1. Experiment 1

The first experimental block shows a clear confusion between CB and IB when CB is not present in the line-up (line-ups 9–12; Tables 1–3). A low number of foils succeeded in distracting the listeners. Over the course of the three blocks there is a movement away from the selection of IB and the selection of a foil halves. Analysis of variance failed to show that the number of correct responses ( $F_{(2,45)}=.608$ ;  $p>.05$ ), false alarms ( $F_{(2,21)}=.237$ ;  $p>.05$ ) or hits ( $F_{(2,21)}=1.699$ ;  $p>.05$ ) changed significantly due to repeated exposure to the target voice.

### 3.2. Experiment 2

The first experimental block shows a stronger confusion between CB and IB than exists for Experiment 1, Block 1 (Table 4). This was expected as the listeners were recognizing CB from memory. The selection of a foil when neither CB or IB is present in the line-up was also greater than in Experiment 1, Block 1.

Over the course of the three blocks, Tables 4–6, three patterns emerge. One, the selection of a foil gradually reduced. Two, the selection of IB when CB is present marginally increases and three, the selection IB when CB is absent decreases. Although there was no increase in the number of hits, there was a decrease in the number of false alarms. Analysis of variance failed to show that the number of correct responses ( $F_{(2,45)}=.202$ ;  $p>.05$ ), false alarms ( $F_{(2,21)}=.381$ ;  $p>.05$ ) or hits ( $F_{(2,21)}=.007$ ;  $p>.05$ ) changed significantly due to repeated exposure to the target voice.

### 3.3. Experiment 4

The target voice in this experiment was NI, the natural voice of the professional imitator. The first experimental block (Table 7) shows a stronger confusion between NI and CB than with IB when CB is present in the line-up. When CB is not present, there is an increase in the degree of confusion between IN and IB. By the second experimental block the selection of a foil has reduced from 15 listeners to just one. There is also a movement to the selection of IN when present over IB. However there is no movement away from the selection CB when both IB and CB are present in the line-up.

Analysis of variance showed a significant change in the number of correct imitations ( $F_{(2,45)}=6.804$ ;  $p=.003$ ), and hits ( $F_{(2,21)}=24.993$ ;  $p<.000$ ) changed significantly due to repeated exposure to the target voice. The change in the number of false alarms, however, was not significant ( $F_{(2,21)}=2.172$ ;  $p>.05$ ).

## 4. CONCLUSIONS

The research reported in [3] demonstrated that repeated exposure to an imitation did not significantly affect listener discrimination.

The research presented here has shown no significant change in listener response solely due to repeated testing (Experiment 2) or due to repeated exposure to the imitated voice (Experiment 1).

In Experiment 4 the task was to identify NI. This voice was not aiming to be another voice as in [3], nor did it have an imitation distracter in the line-up. Although, IB is the same speaker as NI and is, thus, a similar voice. The block 1 responses of the listeners were more broadly spread than for the other experiments; the listeners were not as sure which voice they were to recognize. Repeated exposure increased the listeners' confidence when the target voice was present, yet not when absent from the line-up. Hence the significant change in the number of correct and hit, but not false alarm responses.

The degree of variation in the first block responses can be seen as reflecting the security of voice selection. The weaker this is, the greater the possibility for change and reduction in the impact of the similar voice due to repeated exposure to a target voice. Initial impact of a similar voice, and its ultimate reduction, is probably due to a combination of the degree of similarity of the spectral and articulatory timing aspects (see [4]). This warrants further investigation.

## NOTES

1. The authors are listed in reverse alphabetical order.

## REFERENCES

- [1] Clifford, B. R. 1983. Memory for voices: The feasibility and quality of earwitness evidence. In Lloyd, B. R. and B. R. Clifford (eds.) *Evaluating Witness Evidence*. Chichester, Wiley.
- [2] Schlichting, F. and Sullivan, K. P. H. 1997. The imitated voice — a problem for voice line-ups?. *Forensic Linguistics: The International Journal of Speech, Language and the Law*, 4, 148–165 (Errata 4, 318–319)
- [3] Sullivan, K. P. H. and Schlichting, F. 1998. Increased discrimination — a consequence of repeated exposure to an imitation of a familiar voice? In Branderud, P. and H. Traunmüller (eds.) *Proceedings of FONETIK 98*. Stockholm: Dept. Linguistics, Stockholm University.
- [4] Wretling, P. and Eriksson, A. 1998. Is articulatory timing speaker specific? — Evidence from imitated voices. In Branderud, P. and H. Traunmüller (eds.) *Proceedings of FONETIK 98*. Stockholm: Dept. Linguistics, Stockholm University.