

Julien Plante-Hébert, Victor J. Boucher (2013)  
Université de Montréal

*The effect of familiarity on speaker identification in voice line-ups.*

**Abstract**

**Introduction**

This study examines the validity of the technique of voice line-ups in identifying or discriminating target speakers by focusing on the effect of the familiarity of heard voices. Previous reports have shown that a number of variables have to be controlled in devising voice line-ups containing a specific individual or suspect (e.g. Fernandez Gallardo, 2012; Jessen, 2008; Nolan, 2003; Yarmey, 1995). The present study is the first to examine the central factor of familiarity in terms of a complex index that took into account three related effects: 1) *recency* (i.e. the time of last spoken contact), 2) *duration of spoken contact*, and 3) *frequency of spoken contact*. In the present study, low familiarity reflected the case where, for instance, a listener identified the voice of stranger that was heard a few minutes preceding the experiment (cf. Yarmey, 2001). Very high familiarity was given by cases where a listener heard the voice of his father or twin brother (etc.) with whom he lived for extended periods. The specific purpose was to determine if identifications of heard speakers over a cell phone could achieve a high accuracy (at over 99%) based on a high level of familiarity. We also wanted to know if the length of heard utterances could affect accuracy rates.

**Stimuli and subjects**

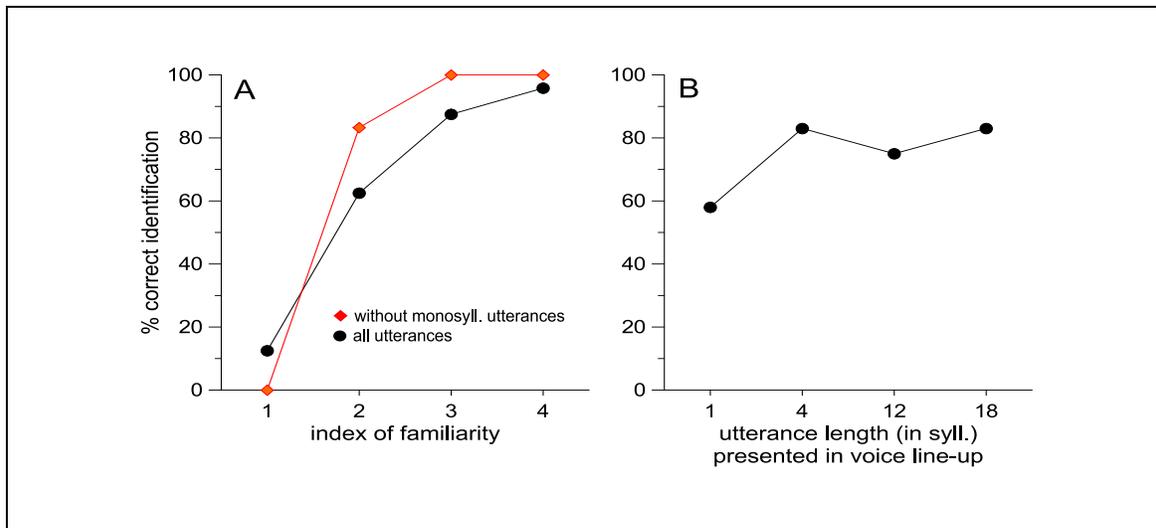
The stimuli were three recorded voice line-ups each constructed using 10 male voices including one of three target voices/speakers. Participants (listeners) in the present study were selected on the basis of their familiarity with the target voices/speakers (e.g. the target speaker was a relative, a life partner, a workmate, an acquaintance, a stranger etc. all categorized in terms of the three familiarity factors noted above). In the design of the voice line-ups, the speaking fundamental frequency (SF0) of all the selected voices did not vary beyond one semitone (i.e., all voices had a similar “modal pitch”). The utterances heard in the line-ups were from native speakers of Québécois French with no strong regional accent. These utterances reflected familiar greetings that varied in four given lengths (utterances contained 1, 4, 12, 18 syllables balanced in terms of their oral and nasal features). Finally, all voices were first recorded in a soundproof room and then filtered in terms of the frequency bands of a typical cellphone. The final stimuli were these filtered recordings.

**Procedure**

The participants were asked to identify a known voice (a father, a brother, an acquaintance, etc.) in a line-up by listening to the recordings through headphones in a quiet room. Participants were specifically instructed to listen to all 10 voices in a line-up before giving an answer and they could playback the recordings and rectify their choice at will.

## Results and discussion

The results of Figure 1 show that high rates of speaker identification are obtained only when the listener is presented with speech from a highly familiar individual. The best rates (100%) were obtained for heard multisyllabic utterances (4 syllables and more) and for levels 3 and 4 of our familiarity index. A level 4 represents a close relative (e.g. father, brother, twin) with whom one has daily interactions over several years. A level 3 reflects close friends and companions with whom one has frequent weekly interactions over extended periods. Overall, these exploratory findings, which reflect the preliminary part of a study involving 30 subjects (to be presented at the Conference), already suggest that, in a carefully constructed voice line-up where speech material is heard at band-pass frequencies of a cellphone, high rates of correct speaker identification can be obtained only with voices that are highly familiar to the listener. Subsequent work using a larger listener base will serve to define functional scales of familiarity.



**Figure 1** A. Correct identification as a function of the index of familiarity. Black line: identification rates for heard utterances of 1-18 syll.; red line, identification rates for utterances of 4-18 syll. B. Identification rates by length of utterance heard in the voice line-up.

Plante-Hébert, J. and Boucher, V. J. (2013), *The effect of familiarity on speaker identification in voice line-ups*. IAFPA annual conference. University of South Florida, July 23rd 2013