

# 143rd ASA Meeting, Pittsburgh, PA

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

## When Sound Effects Are Better Than The Real Thing

**Laurie M. Heller - [laurie\\_heller@brown.edu](mailto:laurie_heller@brown.edu)**

Lauren Wolf

Department of Psychology

Brown University

Providence, RI 02912

Popular version of paper 1pPP10

Presented Monday Afternoon, June 3, 2002

143rd ASA Meeting, Pittsburgh, PA

Sound effect technicians, known as “Foley artists,” have long known that two physically different events can produce similar sounds: the sound of footsteps in the snow can be imitated by squeezing a box of cornstarch. It has also been claimed that sound effects can benefit from exaggerating the sound, but until now this assumption has not been scientifically tested. We chose to study sound effects with the hypothesis that they would provide examples of successful auditory caricatures. This approach to the study of sound identification borrows from the study of the visual recognition of facial caricatures. Caricatures that exaggerate essential facial features have been found to improve recognition. Through the study of auditory caricatures we are interested in discovering what is essential in a sound for its identification.

Our first objective was to generate sounds and identify which ones would be the best candidates for synthesis in the main experiment. Nine examples of traditional Foley sound effects were generated, such as: crackling fire (emulated by twisting cellophane), walking in leaves (running fingers through a box of cornflakes), etc. These examples were generated to emulate nine recordings of the corresponding real events. For example, a recording of twisting cellophane was generated to correspond to a recording of a real fire. Five repeated instances of each of these 18 events were generated, recorded and tested.

A preliminary experiment was run to ensure that we started out with adequate stimuli. Seventeen normal-hearing research volunteers listened over headphones and tried to identify all of the real and Foley auditory events. Listeners were given no information about the sounds, but were told to indicate whatever objects, materials and actions they thought had generated the sounds. Responses were later coded for accuracy in terms of both the material and action. Four of the nine Foley effects were identified more accurately (as their intended effect) than were their real counterparts. Next, each sound was rated as to how realistic it sounded when the intended event (e.g., crackling fire) was indicated to the subject. Although some subjects consistently rated certain Foley effects as more realistic than real events, the real events were rated more realistic, on average, across all listeners.

In the main experiment we attempted to synthesize sounds that would be “more realistic” than the real ones. To generate stimuli, we digitally synthesized new stimuli out of three pairs of sound stimuli that were tested in the preliminary experiment. Our hypothesis was that the real event of “walking in mud” captured the physical action (walking) better than it did the material (mud) whereas the identification data indicated that the Foley “walking in mud” did a better job of conveying the material (mud). Accordingly, we mathematically extracted the acoustic features hypothesized to convey the action (e.g., walking) from the recording of the real event and combined that with the acoustic features that indicated the material (e.g., mud) from a Foley version of “walking in mud.” To extend this technique to more than one sound, an analogous procedure was used to produce synthesized versions of “walking in leaves” and “crushing eggshells.”

A set of 20 new subjects judged the real, Foley, and newly synthesized sounds as comparison pairs. On each trial, listeners heard two sounds in a row, both of which were identified with the same label (e.g., “walking in mud”). Listeners indicated which sound was more realistic. Over 70% of the time (a statistically significant percentage), listeners preferred the synthesized event to either the real or Foley stimulus for the “walking in mud” stimuli. This was also found for the “walking in leaves” stimuli. No significant preference was determined for the “crushing eggshells” stimulus. Future tests will apply this technique to a broader range of sounds.

Ultimately, the results of this project could be useful in later efforts to create audio caricatures through synthesis. This caricature will exaggerate important acoustic features and yet be recognizable as a particular sound event, just as a computer-generated caricature of a face is recognizable as a particular individual.

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