



HAL
open science

Comparing identification of vocal imitations and computational sketches of everyday sounds

Guillaume Lemaître, Olivier Houix, Frédéric Voisin, Nicolas Misdariis, Patrick Susini

► **To cite this version:**

Guillaume Lemaître, Olivier Houix, Frédéric Voisin, Nicolas Misdariis, Patrick Susini. Comparing identification of vocal imitations and computational sketches of everyday sounds. Meeting of the Acoustical Society of America, Nov 2016, Honolulu, United States. pp.3390 - 3390, 10.1121/1.4970854 . hal-01448968

HAL Id: hal-01448968

<https://hal.science/hal-01448968>

Submitted on 29 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparing the identification of vocal imitations and computational sketches of everyday sounds

Guillaume Lemaitre, Olivier Houix, Frédéric Voisin, Nicolas Misdariis, Patrick Susini

Sounds are notably difficult to describe. It is thus not surprising that human speakers often use many imitative vocalizations to communicate about sounds. In practice, vocal imitations of non-speech everyday sounds (e.g. the sound of a car passing by) are very effective: listeners identify sounds better with vocal imitations than with verbal descriptions, despite the fact that vocal imitations are often inaccurate, constrained by the human vocal apparatus. The present study investigated the semantic representations evoked by vocal imitations by experimentally quantifying how well listeners could match sounds to category labels. It compared two different types of sounds: human vocal imitations, and computational auditory sketches (created by algorithmic computations), both based on easily identifiable sounds (sounds of human actions and manufactured products). The results show that performance with the best vocal imitations was similar to the best auditory sketches for most categories of sounds. More detailed analyses showed that the acoustic distance between vocal imitations and referent sounds is not sufficient to account for such performance. They suggested that instead of reproducing the acoustic properties of the referent sound as accurately as vocally possible, vocal imitations focus on a few important features dependent on each particular sound category.