FP7-ICT-2013-C FET-Future Emerging Technologies-618067

SkAT-VG: Sketching Audio Technologies using Vocalizations and Gestures

D2.2.1
Explorative collection of imitated sounds

First Author	Pétur Helgason
Responsible Partner	KTH
Status-Version:	Draft-1.0
Date:	March 18, 2015
EC Distribution:	Consortium
Project Number:	618067
Project Title:	Sketching Audio Technologies using Vocalizations and Gestures

Title of Deliverable:	Explorative collection of imitated sounds
Date of delivery to the EC:	23/05/2014

Workpackage responsible for the Deliverable	WP2
Editor(s):	Davide A. Mauro
Contributor(s):	Anders Friberg, Sofia Strömbergsson, Sten Ternström, Davide A. Mauro, Davide Rocchesso, Guillaume Lemaitre
Reviewer(s):	Davide A. Mauro
Approved by:	All Partners

Abstract	The scope of the current deliverable is to present the results of task T2.1: Case studies will be collected from commercial recordings. Their quality will probably not be adequate, but they will be inspirational and reveal what skilled imitators can do.
Keyword List:	sound imitation

Document Revision History

Deliverable D2.2.1 was due on april 30, 2014. The date of delivery was delayed due to illness of the principal author and to some difficulties the project experienced in setting the teams up in its early months.

Further versions of this document may be published until the end of work package WP2.

Version	Date	Description	Author
First draft	19/05/2014	Import from .doc template	DAM
v0.2	20/05/2014	Contributions and updates	ROC
v0.3	21/05/2014	Minor corrections	SOS
v0.4	23/05/2014	Adding list of imitations from previous studies	GLM

Table of Contents

Index of Figures

List of Acronyms and Abbreviations

DoW Description of Work

EC European Commission

PM Person Months

WP Work Package

1 Introduction

This deliverable contains four collections of recordings:

1. A Vocalizations showcase, focusing on which kinds of sounds people can produce in imitations and how these sounds are made.
2. A Sound origins showcase, focusing on illustrating imitations (or impressions) of sound action primitives. The defined classes of sound action primitives relating to basic mechanical interactions match closely those used by [LDSA11] (which focused on solids, liquids and gasses). However, complementary to this, impressions of animals and engines are also given, since it seems clear that imitators use different aspects of their sound production mechanism for imitating basic mechanical interactions on the one hand and animal and engine noises on the other.
3. The collection of sounds from the book Mouth Sounds [New04] compiled by IUAV.
4. Imitations and referent sounds used in [LR14, LDSA11].

As an appendix, we deliver also Pétur Helgasons paper to Fonetik 2014.

2 Vocalizations showcase

The Vocalizations showcase is a collection of short video clips (with sound), which serve to demonstrate the ways in which people can use their sound production apparatus to do imitations. The video files have been partly collected on-line, and so quality and format will differ. To a large extent these different ways of producing sounds are discussed in the Fonetik 2014 paper [Hel14], included as an appendix. The recordings are indexed in the file `Vocalizations.htm` (see Table 1) which also offers a convenient browsing interface.

cccccc

Periodicity	Source type	Initiation	Source origin	Airstream	Example
Periodic	Myoelastic	Pulmonic	Modal voice	Egressive	Cow mooing M01
				Ingressive	No example available
			Falsetto voice	Egressive	Cat meowing F01
				Ingressive	Donkey braying M01
			Creaky voice	Egressive	To be indexed
				Ingressive	Frog croaking FN (sound only)
			Breathy voice	Egressive	To be indexed
				Ingressive	To be indexed
			Aryepiglottic vibration/trill	Egressive	To be indexed
				Ingressive	Not possible
			Ventricular vibration/trill	Egressive	Lion roaring F01
				Ingressive	Not possible
			Velic trill	Egressive	Not possible
				Ingressive	Pig snorting M01
			Uvular trill	Egressive	Crow cawing M01
				Ingressive	No example available
			Apico-alveolar trill	Egressive	Motorbike F01
				Ingressive	Not possible
			Bilabial trill	Egressive	Lawn mower M01
				Ingressive	Not possible
			Bilabial vibration	Egressive	To be indexed
				Ingressive	Not possible
		Glottalic	Producing a myoelastic source with a glottalic airstream is difficult. Apico-alveolar and bilabial myoelastic source are possible to make but they cannot be maintained for more than a few cycles (4-8, depending on the individual).		
		Velaric	Producing a myoelastic source with a velaric airstream is difficult. A bilabial myoelastic source is possible to make but it cannot be maintained for more than a few cycles (2-4, depending on the individual).		
	Whistled	Pulmonic	Labial	Egressive	To be indexed

				Ingressive	No example available
			Alveolo-palatal	Egressive	Bird & cricket N01
				Ingressive	Not possible?
		Glottalic	Labial	Egressive	No example available
				Ingressive	No example available
			Alveolo-palatal	Egressive	No example available
				Ingressive	No example available
		Velaric	Labial	Egressive	Drop of water N02
				Ingressive	No example available
Aperiodic	Turbulent	Pulmonic	Glottal	Egressive	To be indexed
				Ingressive	No example available
			Pharyngeal	Egressive	To be indexed
				Ingressive	No example available
			Uvular	Egressive	To be indexed
				Ingressive	No example available
			Velar	Egressive	To be indexed
				Ingressive	To be indexed
			Pre-velar	Egressive	To be indexed
				Ingressive	To be indexed
			Palatal	Egressive	To be indexed
				Ingressive	No example available
			Post-alveolar flat	Egressive	To be indexed
				Ingressive	To be indexed
			Alveolar flat	Egressive	To be indexed
				Ingressive	To be indexed
			Dental flat	Egressive	To be indexed
				Ingressive	To be indexed
			Post-alveolar grooved	Egressive	To be indexed
				Ingressive	Not possible
			Alveolar grooved	Egressive	To be indexed
				Ingressive	Not possible
			Dental grooved	Egressive	To be indexed
				Ingressive	Not possible
			Labiodental	Egressive	To be indexed
				Ingressive	No example available
			Bilabial	Egressive	To be indexed
				Ingressive	No example available
			Velar lateral	Egressive	To be indexed
				Ingressive	No example available
			Pre-velar lateral	Egressive	To be indexed
				Ingressive	No example available
			Palatal lateral	Egressive	To be indexed
				Ingressive	No example available
			Post-alveolar lateral	Egressive	To be indexed
				Ingressive	No example available
			Alveolar lateral	Egressive	To be indexed
				Ingressive	No example available
			Dental lateral	Egressive	To be indexed
				Ingressive	No example available
		Glottalic	Although glottalic initiation can be used for all the source places that exist for pulmonic initiation (excepting glottal), our data contain no examples of glottalic initiation for turbulent sources.		
		Velaric	Palatal	Egressive	No example available
				Ingressive	To be indexed
			Post-alveolar (retroflex)	Egressive	No example available
				Ingressive	To be indexed
			Alveolar flat	Egressive	No example available
				Ingressive	To be indexed
			Alveolar lateral	Egressive	No example available
				Ingressive	Camera click F01
			Dental flat	Egressive	No example available
				Ingressive	To be indexed
			Labiodental	Egressive	No example available
				Ingressive	No example available
			Bilabial	Egressive	No example available
				Ingressive	No example available
	Percussive	Dorsal	Sublaminal	Not appl.	To be indexed
			Lamino-dental	Not appl.	To be indexed
		Mandibular	Dental	Not appl.	To be indexed

Table 1: Table describing the vocalizations recorded.

3 Sound origins showcase

The Sound origins showcase features a collection of recently recorded sound files with imitations.[1] These are indexed with regard to sound action primitives and other qualities pertaining to the origin of the sound: animals and engines. The focus is on providing several different

[1]Simultaneous video recordings were also made, but are not part of this deliverable.

examples of each class/type of sound.

Each sound clip in the Sound origins showcase contains a whole instance of a single imitation (impression) and therefore the durations of the sound files differ (some may be 300 ms while others go on for seconds). The files are stereo, with different gains on the channels, i.e. the sound may be louder on one channel. Each file is named according to the following code:

$$< Sound - ID > \_ < Speaker - ID > \_ < Trial - ID > \quad (1)$$

The $< Sound - ID >$ includes the class of sound, a subclass, the type of sound and, in some cases, a subtype.

The $< Speaker - ID >$ includes only the speaker initials.

The $< Trial - ID >$ identifies different repetitions of the same imitation by a speaker.

Below is an example of what the index (i.e. file name) for a single sound file looks like:
`SS03B_MB_03.wav`

- S = The class of sound is Sound-action related
- S = The subclass is sound emanating in solids
- 03 = The type of sound is scraping
- B = The subtype is scraping sand paper against wood
- MB = The subject's initials
- 03 = This is the third repetition of this imitation provided by the subject.

In some cases there are no subclasses or subtypes, in which case fillers are used to ensure that the onset position of the codes is the same in all files. For example, for engines we have not defined a subclass so all engine sounds start with EE, in which the second E serves as a filler.

3.1 Recordings

3.1.1 Speakers

Three speakers were recruited for participation. Descriptive data for these speakers are summarized in Table 2. All listeners reported normal hearing, and no previous experience of vocal imitation, beat boxing or similar.

Speaker ID	Gender	Age (years)	Residence History
F01	Female	26	Sweden since birth
F02	Female	44	Sweden since birth; 3 years in Norway as an adult
M01	Male	29	Sweden since birth

Table 2: Descriptions of participating speakers.

3.1.2 Recording setup

Recordings were performed in a sound-proof studio. Audio was recorded via two head-mounted microphones, set at different gains, to 44.1 kHz stereo files using Audacity or Adobe Audition. The microphones were both condenser type, a DPA 4066 omnidirectional, through a Grace M101 microphone pre-amp, and an AKG C520 cardioid, feeding a MOTU Traveler digital audio interface to a MacBook Pro computer. Video was recorded using a GoPro H3+ camera but is not part of this deliverable.

3.1.3 Recording protocol

A recording protocol was designed, based on the Sound Action Primitives (SAPs) as described in [LDSA11], but adapted to Swedish. Hence, sounds originating from Solids, Liquids and Gasses were represented. The protocol was expanded to include Engine sounds and Animal sounds, in the hope that these would elicit more complex articulatory actions from the speakers, e.g. more voiced sounds. The recording script was updated between recordings, and some sound types are therefore not produced by all speakers.

3.2 Annotation and data processing

The recordings were annotated in ELAN (available from `http://tla.mpi.nl/tools/tla-tools/elan/`). The original recording (audio) was opened in ELAN, together with the annotation template `SkAT_template.etf`, in which a restricted set of annotation labels are defined. Vocalizations were annotated with regards to these labels, thereby indexing the type of vocalization and when it occurred in the original recording. The format of the sound type labels XXNNY, where XX is a code indexing the main sound class (`SS = SAP: Solids; SL = SAP: Liquids; SG = SAP: Gases; AA = Animals; EE = Engines`), NN refers to the subclass within this category, and Y represents the subtype of this sound. For a complete list of sound labels, see section 3.3.1. During annotation, no evaluation was made of whether the imitation was successful or not. Hence, all efforts at imitating a specific sound were annotated, regardless of whether the speaker (or the experimenter) was satisfied or not. Vocalizations turning into laughter were annotated when possible, but in the event of further processing and analysis, these vocalizations deserve cautious treatment. The ELAN annotations were exported into TextGrid-format, to allow automatic extraction of the vocalizations segments in Praat. In this step, the sound clips were automatically indexed with information about speaker ID and trial index. Hence, the format of the sound clip filenames is $< Sound - ID > \_ < Speaker - ID > \_ < Trial - ID >$ where $< Sound - ID >$ refers to the sound type label (as listed in section 3.3.1), $< Speaker - ID >$ is the ID of the speaker, and $< Trial - ID >$ identifies different repetitions of the same imitation by a speaker.

3.3 Data structure

All sound clips (vocalizations) are contained in the folder `SkATrecs/`. The complete set of clips is listed and described in the TAB-separated document `SkATrecsDesc.txt`. The folder `HTML/` contains html-files that allows structured overview of the dataset. The index

file `SkATrecs_expl.html` displays the Table 3 of hyperlinks (to the html-files in the folder `HTML/`):

Spkr	CATEGORIES				
	Animal_sounds	Engine_sounds	SAP_Solids	SAP_Gasses	SAP_Liquids
F01	F01:Animal_sounds	F01:Engine_sounds	F01:SAP_Solids	F01:SAP_Gasses	F01:SAP_Liquids
F02	F02:Animal_sounds	F02:Engine_sounds	F02:SAP_Solids	F02:SAP_Gasses	F02:SAP_Liquids
M01	M01:Animal_sounds	M01:Engine_sounds	M01:SAP_Solids	M01:SAP_Gasses	M01:SAP_Liquids

Table 3: Table of the index page to access contents.

From this table, different sets of sound files are easily accessible, e.g. all vocalizations from speaker F01, all Solid sounds from all speakers, or all Solid sounds produced by speaker F01.

3.3.1 Sound codes

Code	Class	Class label	Subclass	Subtype
AA01A	Animal	AA	cat	small cat
AA01B	Animal	AA	cat	big cat
AA02A	Animal	AA	cat	cat
AA03A	Animal	AA	dog	small dog
AA03B	Animal	AA	dog	big dog
AA04A	Animal	AA	dog	dog
AA05A	Animal	AA		cow
AA06A	Animal	AA	horse	horse (gngg)
AA06B	Animal	AA		horse (frust)
AA07A	Animal	AA		horse (frust)
AA08A	Animal	AA		horse (clammering hooves)
AA09A	Animal	AA		elephant
AA10A	Animal	AA		pig (snort)
AA11A	Animal	AA		lion
AA12A	Animal	AA		crow
AA13A	Animal	AA		small bird (kvitter)
AA14A	Animal	AA		seagull
AA15A	Animal	AA		donkey
AA16A	Animal	AA		frog
EE01A	Engine	EE	car	an ordinary car
EE01B	Engine	EE	car	an ordinary car, stationary
EE01C	Engine	EE	car	an ordinary car, accelerating
EE02A	Engine	EE		racing car
EE03A	Engine	EE	truck	truck
EE03B	Engine	EE	truck	truck, stationary
EE03C	Engine	EE	truck	truck, accelerating
EE04A	Engine	EE		moped/vespa
EE05A	Engine	EE		motor bike
EE06A	Engine	EE		vespa
EE07A	Engine	EE		electric razor

EE08A	Engine	EE		small robot arm
EE09A	Engine	EE		lawn mower
EE10A	Engine	EE		leaf blower
EE11A	Engine	EE		helicopter
EE12A	Engine	EE		camera click
SG01A	SAP: Gases	SG	gun shot	rifle
SG01B	SAP: Gases	SG	gun shot	gun
SG01C	SAP: Gases	SG	gun shot	air gun
SG02A	SAP: Gases	SG	blasting	explosion (firework)
SG03A	SAP: Gases	SG	whipping	whip
SG04A	SAP: Gases	SG	puffs	bicycle pump
SG05A	SAP: Gases	SG	blowing	ventilation drum
SG05B	SAP: Gases	SG	blowing	pipe leak
SG06A	SAP: Gases	SG	whirling	small fan (hand held)
SS01A	SAP: Solids	SS	scraping	scraping two stones
SS01B	SAP: Solids	SS	scraping	sandpaper against wood
SS01C	SAP: Solids	SS	scraping	hand-sawing through wood
SS02A	SAP: Solids	SS	squeaking	squeaking glass
SS02B	SAP: Solids	SS	squeaking	an old teeter board
SS03A	SAP: Solids	SS	tearing	tearing some paper
SS04A	SAP: Solids	SS	crumpling	crumpling some paper
SS05A	SAP: Solids	SS	slamming	slamming hand on a board
SS06A	SAP: Solids	SS	tapping	tapping finger on a board
SS06B	SAP: Solids	SS	tapping	tapping with a pipe
SL01A	SAP: Liquids	SL	gushing	gushing water
SL02A	SAP: Liquids	SL	trickling	trickling water
SL03A	SAP: Liquids	SL	spattering	spattering water
SL04A	SAP: Liquids	SL	sloshing	sloshing water
SL05A	SAP: Liquids	SL	dripping	water dripping in an empty sink
SL05B	SAP: Liquids	SL	dripping	water dripping in a sink filled with water

Table 4: Table describing the soundcodes produced.

4 Mouth Sounds

A collection of 152 audio segments have been manually extracted from the sound examples of the companion CD of the book Mouth Sounds [New04]. Each of these audio segments is 500ms-long. The audio segments are described in Table 5. l The column "Decomposition" refers to the automatic separation of the sound excerpt into its sinusoidal, noise, and transient components. This separation has been performed by Hélène Lachambre using a Genesis proprietary software tool. An example of decomposition is illustrated in Figure 1.

Filename	Mechanism	Description	Decomposition	Reference
track03ex01.wav	? (whistle?)			Mouth Sounds, ch. 2 Falsetto
track04ex01.wav	Ploit	Prototypical	X	Mouth Sounds, ch. 2 Ploit
track04ex02.wav	Pop	Pop variant	X	Mouth Sounds, ch. 2 Ploit

track04ex03.wav	Pop	Pop variant	X	Mouth Sounds, ch. 2 Ploit
track04ex04.wav	Pop	Pop variant	X	Mouth Sounds, ch. 2 Ploit
track05ex01.wav	Palate grind	Prototypical	X	Mouth Sounds, ch. 2 Palate Grind
track05ex02.wav	Palate grind	Prototypical	X	Mouth Sounds, ch. 2 Palate Grind
track05ex03.wav	Palate grind	Palate Grind (exhale)	X	Mouth Sounds, ch. 2 Palate Grind
track05ex04.wav	Palate grind	Prototypical	X	Mouth Sounds, ch. 2 Palate Grind
track06ex01.wav	Glottal Fry	Prototypical	X	Mouth Sounds, ch. 2 Glottal Fry
track06ex02.wav	Glottal Fry	Door		Mouth Sounds, ch. 2 Glottal Fry
track06ex03.wav	Glottal Fry	Glottal fry inhaling	X	Mouth Sounds, ch. 2 Glottal Fry
track06ex04.wav	Glottal Fry	Gun		Mouth Sounds, ch. 2 Glottal Fry
track07ex01.wav	?			Mouth Sounds, ch. 3 Champagne Pop
track07ex02.wav	?			Mouth Sounds, ch. 3 Champagne Pop
track07ex03.wav	Pop	Pop	X	Mouth Sounds, ch. 3 Champagne Pop
track08ex01.wav	Tongue Flop	Horse Trot (?)	X	Mouth Sounds, ch. 3 Tongue Flop
track09ex01.wav	Suction Pop	Prototypical	X	Mouth Sounds, ch. 3 Suction Pop
track09ex02.wav	Palate Grind (?)			Mouth Sounds, ch. 3 Suction Pop
track10ex01.wav	Slap Pop (Ploit variant)	Prototypical	X	Mouth Sounds, ch. 3 Slap Pop
track11ex01.wav	Ping Pong Pop	Prototypical	X	Mouth Sounds, ch. 3 Ping Pong Pop
track12ex01.wav	Triple click	Prototypical	X	Mouth Sounds, ch. 3 Triple Click
track13ex01.wav	Finger Whip	Prototypical		Mouth Sounds, ch. 3 Finger Whip
track14ex01.wav	Honk	Prototypical		Mouth Sounds, ch. 3 Basic Honk
track14ex02.wav	Honk	Bicycle horn	X	Mouth Sounds, ch. 3 Basic Honk
track14ex03.wav	Honk	Air horn	X	Mouth Sounds, ch. 3 Basic Honk
track14ex04.wav	Honk	Goose honk	X	Mouth Sounds, ch. 3 Basic Honk
track15ex01.wav	?	Prototypical		Mouth Sounds, ch. 3 Ocean Liner Blast
track15ex02.wav	?		X	Mouth Sounds, ch. 3 Ocean Liner Blast
track15ex03.wav	?		X	Mouth Sounds, ch. 3 Ocean Liner Blast
track15ex04.wav	?	Gun	X	Mouth Sounds, ch. 3 Ocean Liner Blast
track16ex01.wav	?	Prototypical	X	Mouth Sounds, ch. 3 Tugboat Blast
track17ex01.wav	?	Whistle	X	
track18ex01.wav	?		X	Mouth Sounds, ch. 4 Dog Bark
track18ex02.wav	?		X	Mouth Sounds, ch. 4 Dog Bark
track19ex01.wav	?		X	Mouth Sounds, ch. 4 Cat Meow
track19ex02.wav	?		X	Mouth Sounds, ch. 4 Cat Meow
track20ex01.wav	?		X	Mouth Sounds, ch. 4 Donkey Bray
track22ex01.wav	?			Mouth Sounds, ch. 4 Pig Squeal
track22ex02.wav	?		X	Mouth Sounds, ch. 4 Pig Snort
track22ex03.wav	?		X	Mouth Sounds, ch. 4 Pig Snort
track23ex01.wav	Falsetto		X	Mouth Sounds, ch. 4 Chicken Cluck
track23ex02.wav	?		X	Mouth Sounds, ch. 4 Chicken Flutters
track23ex03.wav	?		X	Mouth Sounds, ch. 4 Chicken Flutters
track23ex04.wav	?		X	Mouth Sounds, ch. 4 Pigeon Coo
track25ex01.wav	Whistle		X	Mouth Sounds, ch. 4 Bird Tweets
track25ex02.wav	Whistle			Mouth Sounds, ch. 4 Bird Tweets
track25ex03.wav	Inhale			Mouth Sounds, ch. 4 Bird Squawks
track26ex01.wav	Falsetto			Mouth Sounds, ch. 4 Peacock Wail
track26ex02.wav	Glottal stops			Mouth Sounds, ch. 4 Generic Jungle Bird
track27ex01.wav	Pucker Whistle			Mouth Sounds, ch. 4 Cricket Chirp
track28ex01.wav	Buzzing of the lips			Mouth Sounds, ch. 4 Fly Buzz
track28ex02.wav	Falsetto	Mosquito		Mouth Sounds, ch. 4 Fly Buzz
track28ex03.wav	?	Whip		Mouth Sounds, ch. 4 Fly Buzz
track30ex01.wav	Breath in and out	Chimpanzees		Mouth Sounds, ch. 4 Monkeys and such
track30ex02.wav	Breath in and out	Chimpanzees		Mouth Sounds, ch. 4 Monkeys and such
track31ex01.wav	?	Elephant		Mouth Sounds, ch. 4 Elephant Trumpet
track32ex01.wav	Inhale glottal fry	Frog		Mouth Sounds, ch. 4 Frog Croak
track32ex02.wav	Inhale glottal fry	Frog variant		Mouth Sounds, ch. 4 Frog Croak
track32ex03.wav	Inhale glottal fry	Frog variant		Mouth Sounds, ch. 4 Frog Croak
track32ex04.wav	Inhale glottal fry	Frog variant		Mouth Sounds, ch. 4 Frog Croak
track32ex05.wav	Inhale glottal fry	Frog variant		Mouth Sounds, ch. 4 Frog Croak
track32ex06.wav	Inhale glottal fry	Toad		Mouth Sounds, ch. 4 Frog Croak
track33ex01.wav	?	?		Mouth Sounds, ch. 5 Duck Quack
track35ex01.wav	?	Turkey		Mouth Sounds, ch. 4 Turkey Gaa-Laa
track36ex01.wav	Fingerless Whistle	Seagull		Mouth Sounds, ch. 4 The Seagull
track36ex02.wav	Fingerless Whistle	Seagull		Mouth Sounds, ch. 4 The Seagull
track36ex03.wav	Fingerless Whistle	Whale		Mouth Sounds, ch. 4 The Seagull
track36ex04.wav	?	?		Mouth Sounds, ch. 4 The Seagull
track36ex05.wav	Finger Whisle	Loon Wail		Mouth Sounds, ch. 4 The Loon
track37ex01.wav	Ploit	Fish		Mouth Sounds, ch. 4 The Fish
track37ex02.wav	Ploit	Fish bubble effect		Mouth Sounds, ch. 4 The Fish
track37ex03.wav	Duck squeal	Dolphin		Mouth Sounds, ch. 4 The Dolphin
track37ex04.wav	Glottal fry	Dolphin		Mouth Sounds, ch. 4 The Dolphin
track37ex05.wav	Ventriloquist effect exhale	Dolphin		Mouth Sounds, ch. 4 The Dolphin
track39ex01.wav	Invisible whistle	Buzzer		Mouth Sounds, ch. 5 Invisible Whistle
track39ex02.wav	Invisible whistle	Touch tones telephone		Mouth Sounds, ch. 5 Invisible Whistle
track40ex01.wav	Fingerless Whistle			Mouth Sounds, ch. 5 Fingerless Whistle
track42ex01.wav	Double Whistle			Mouth Sounds, ch. 5 Double Whistle
track43ex01.wav	Nose Whistle	Paper nose whistle		Mouth Sounds, ch. 5 Nose Whistle
track44ex01.wav	?	?		Mouth Sounds, ch. 6 Musical Instruments
track46ex01.wav	?	Classic guitar string		Mouth Sounds, ch. 6 Classical Guitar
track48ex01.wav	?	Bass guitar string		Mouth Sounds, ch. 6 Jazz Bass
track49ex01.wav	?	Bagpipes		Mouth Sounds, ch. 6 Throat Bagpipes
track50ex01.wav	?	Trumpet		Mouth Sounds, ch. 6 Trumpet
track51ex01.wav	?	Muted Trumpet		Mouth Sounds, ch. 6 Muted Trumpet
track52ex01.wav	?	French Horn		Mouth Sounds, ch. 6 French Horn
track53ex01.wav	?	Trombone		Mouth Sounds, ch. 6 The Trombone
track53ex02.wav	?	Cymbals		Mouth Sounds, ch. 6 The Trombone
track54ex01.wav	?	Saxophone		Mouth Sounds, ch. 6 The Saxophone
track54ex02.wav	?	Drums		Mouth Sounds, ch. 6 The Saxophone
track54ex03.wav	?	Dirty Saxophone		Mouth Sounds, ch. 6 The Saxophone
track55ex01.wav	?	Kick		Mouth Sounds, ch. 6 The Drums
track55ex02.wav	?	Marching Snare		Mouth Sounds, ch. 6 The Drums
track55ex03.wav	?	Jazz Snare		Mouth Sounds, ch. 6 The Drums

track55ex04.wav	?	Jazz Snare	Mouth Sounds, ch. 6 The Drums
track55ex05.wav	?	Tom tom	Mouth Sounds, ch. 6 The Drums
track55ex06.wav	?	Tom tom	Mouth Sounds, ch. 6 The Drums
track55ex07.wav	?	Cymbal	Mouth Sounds, ch. 6 The Drums
track55ex08.wav	?	Hi-Hat	Mouth Sounds, ch. 6 The Drums
track56ex01.wav	?	Human beat box	Mouth Sounds, ch. 6 The Drums
track57ex01.wav	?	Steel Drum	Mouth Sounds, ch. 6 Steel Drums
track58ex01.wav	?	Hawaiian nose hum	Mouth Sounds, ch. 6 Hawaiian Nose Hum
track60ex01.wav	?	Saw?	Mouth Sounds, ch. 7 Miscellaneous Sounds
track61ex01.wav	?	Clunker	Mouth Sounds, ch. 7 The Clunker
track61ex02.wav	?		Mouth Sounds, ch. 7 The Clunker
track61ex03.wav	?		Mouth Sounds, ch. 7 The Clunker
track61ex04.wav	Loon	Spinning tires	Mouth Sounds, ch. 7 Spinning Tires
track61ex05.wav	Loon	Spinning tires	Mouth Sounds, ch. 7 Spinning Tires
track62ex01.wav	?	Car Horn	Mouth Sounds, ch. 7 The Sports Car
track62ex02.wav	?	Car	Mouth Sounds, ch. 7 The Sports Car
track62ex03.wav	?	Car	Mouth Sounds, ch. 7 The Sports Car
track62ex04.wav	Palate Grind (?)	Car	Mouth Sounds, ch. 7 The Crash
track63ex01.wav	Falsetto wail	Siren	Mouth Sounds, ch. 7 Sirens and Alarms
track65ex01.wav	sport car + doppler	Cessna	Mouth Sounds, ch. 7 The Flyby
track65ex02.wav	?	Helicopter	Mouth Sounds, ch. 7 The Helicopter
track66ex01.wav	?	Gunshot	Mouth Sounds, ch. 7 Small-arms Fire
track67ex01.wav	Pucker Whistle	Space Effect	Mouth Sounds, ch. 7 Hum and Whistle
track67ex02.wav	?	Lasergun	Mouth Sounds, ch. 7 The Lasergun
track67ex03.wav	?	light saber	Mouth Sounds, ch. 7 The Light Saber
track68ex01.wav	Ploit	?	Mouth Sounds, ch. 7 The Water Drip
track68ex02.wav	Ploit	?	Mouth Sounds, ch. 7 The Water Drip
track68ex03.wav	?	Fire extinguisher	Mouth Sounds, ch. 7 Fire Extinguisher
track69ex01.wav	Invisible whistle		Mouth Sounds, ch. 7 Touch Dialing
track69ex02.wav	?		Mouth Sounds, ch. 7 Touch Dialing
track69ex03.wav	?		Mouth Sounds, ch. 7 Touch Dialing
track69ex04.wav	?		Mouth Sounds, ch. 7 Cellphone ring
track70ex01.wav	Palate grind		Mouth Sounds, ch. 7 Door Knocks
track70ex02.wav	Palate grind		Mouth Sounds, ch. 7 Door Knocks
track70ex03.wav	?	?	Mouth Sounds, ch. 7 Door Knocks
track71ex01.wav	?	?	Mouth Sounds, ch. 7 The Tennis Match
track72ex01.wav	?	?	Mouth Sounds, ch. 8 Rude Noises
track72ex02.wav	?	?	Mouth Sounds, ch. 8 Rude Noises
track73ex01.wav	?	Fart	Mouth Sounds, ch. 8 Raspberry
track74ex01.wav	"Inhale ""oreo"""		Mouth Sounds, ch. 8 Stomach growl
track74ex02.wav	?		Mouth Sounds, ch. 8 The Nose Blow
track74ex03.wav	?		Mouth Sounds, ch. 8 The Nose Blow
track75ex01.wav	?		Mouth Sounds, ch. 8 The Splat
track76ex01.wav	?		Mouth Sounds, ch. 8 Gummy cheeks
track76ex02.wav	?		Mouth Sounds, ch. 8 The Boing
track76ex03.wav	?		Mouth Sounds, ch. 8 The Smock
track76ex04.wav	?		Mouth Sounds, ch. 8 The Mud Trudge
track82ex01.wav	?		Mouth Sounds, ch. 8 The Vent Voice
track82ex02.wav	?		Mouth Sounds, ch. 8 The Vent Voice
track88ex01.wav	?		Mouth Sounds, ch. 8 Meep! Meep!
track89ex01.wav	?		Knuckles the cat
track89ex02.wav	?		Knuckles the cat
track89ex03.wav	?		Knuckles the cat
track89ex04.wav	?		Knuckles the cat
track89ex05.wav	?		Knuckles the cat
track89ex06.wav	?		Knuckles the cat
track89ex07.wav	?		Knuckles the cat

Table 5: Categorization of MouthSounds excerpts.

5 Referent sounds and imitations used in previous studies

The authors of two previous studies have shared the referent sounds and the collection of imitations they collected ([LDSA11, LR14].

5.1 Lemaitre et al., 2011

These sounds were used in "Vocal Imitations and the Identification of Sound Events" by Guillaume Lemaitre, Arnaud Dessein, Patrick Susin & Karine Aura, Ecological Psychology, 2011. There are two folders. The first folder ("Sons originaux") contains the 12 referent sounds. They are described in Table 6.

The second folder ("ImitationsClean") contains 72 imitations of these 12 referent sounds. There were six speakers (three male, three female). The name of the file contains the following information: Gender and number of speakers, referent sound For instance M1_E1 is the first

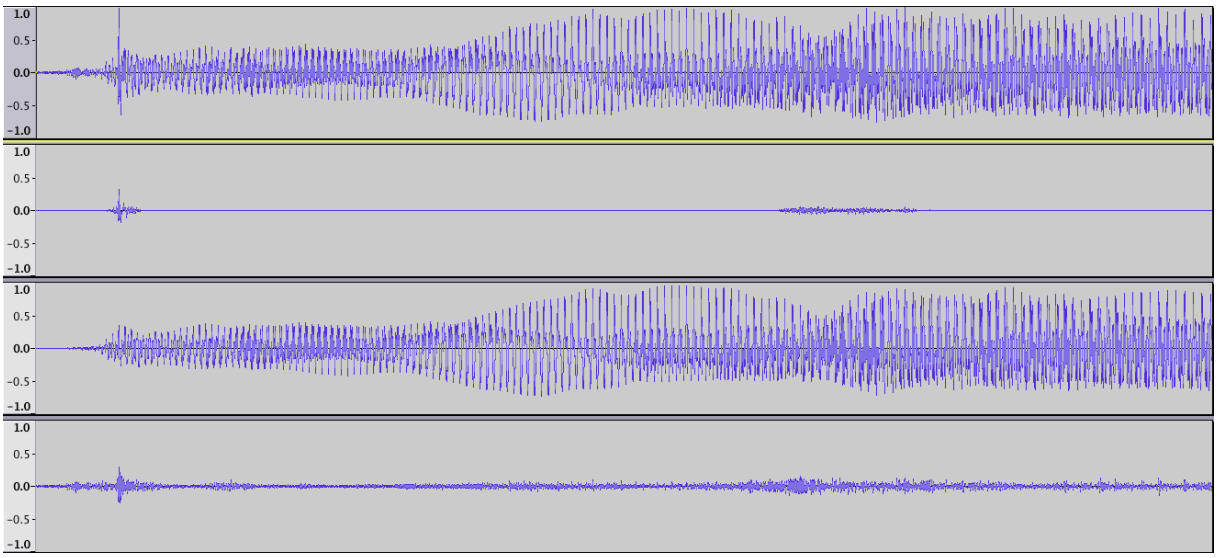

Figure 1: Example of sound decomposition. From upper to lower: original sound, transient, tonal, residual noise.

male imitator imitating the sound E1.

5.2 Lemaitre and Rocchesso, 2014

These sounds were used in "Vocal imitations communicate sounds more effectively than verbalizations" by Guillaume Lemaitre and Davide Rocchesso, Journal of the Acoustical Society of America. 2014. There are four folders corresponding to sections II to IV of the article.

The first folder contains 58 referent sounds. These sounds are listed in Tables 7 and 8. They all were custom-made recordings or synthesis.

The second folder ("SelectionOfReferentSounds") contains a selection of 36 referent sounds, selected from the 58 previously listed sounds.

The third folder ("Imitations") contains 288 imitations. Eight persons (four male, four female) imitated the 36 referent sounds. The 36 imitations of each speakers are organized in the eight folders described in Table 9.

Finally, the fourth folder ("SelectionOfImitations") contains a selection of 108 imitations (the most effective). First part of the file name describes the referent sound. Second part describes the selected speaker. For instance, "C_ChitarraSemi-Acustic_2_Speaker10.wav" is the imitation of the sound "C_ChitarraSemi-Acustic_2.wav" by speaker 10.

Filename	Description	Origin
E1.wav	A tumble dryer	Commercial database
E2.wav	A mixer	Commercial database
E3.wav	A blender	Commercial database
G1.wav	Lighting a match	Commercial database
G2.wav	Gaz from a stove	Commercial database
G3.wav	A spray	Commercial database
L1.wav	Coffee machine	Commercial database
L2.wav	Water dripping	Commercial database
L3.wav	A faucet	Commercial database
S1.wav	Cutting bread	Commercial database
S2.wav	Closing a jar	Commercial database
S3.wav	Cuttin carrots	Commercial database

Table 6: The nine referent sounds used in [LDSA11]

Filename	Description
C_Accendino_1_1.wav	Cigarette lighter
C_Accendino_1_3.wav	Cigarette lighter
C_Accendino_2_2.wav	Cigarette lighter
C_Accendino_3_1.wav	Cigarette lighter
C_ChitarraSemi-Acustic_1.wav	Electric guitar, unplugged
C_ChitarraSemi-Acustic_2.wav	Electric guitar, unplugged
C_ChitarraSemi-Acustic_3.wav	Electric guitar, unplugged
C_ChitarraSemi-Acustic_4.wav	Electric guitar, unplugged
C_ColtelliPiatto_1.wav	Knifes rubbed on a ceramic plate
C_ColtelliPiatto_10.wav	Knifes rubbed on a ceramic plate
C_ColtelliPiatto_3.wav	Knifes rubbed on a ceramic plate
C_ColtelliPiatto_7.wav	Knifes rubbed on a ceramic plate
C_Monete_1.wav	Coins
C_Monete_2.wav	Coins
C_Monete_3.wav	Coins
C_Monete_4.wav	Coins
E_Blowing2_64.wav	Blowing
E_Bouncing_Wood_Big_081.wav	A wooden pipe bouncing
E_Crumpling2_75.wav	Crumpling a piece of paper
E_Dribbling5.wav	Dribbling
E_Dripping5_54.wav	Dripping
E_Hitting_Wood_Big_090.wav	A wooden pipe hitting a surface
E_Leaking.wav	Leaking
E_Puffing3.wav	Puffing
E_Rolling_Glass_Big_077.wav	A glass pipe rolling on a surface
E_Scraping_Metal_Big_072.wav	Scraping a sheet of metal
E_Sloshing2_62.wav	Sloshing
E_Splattering4_75.wav	Splattering
E_Whipping1_58.wav	Whipping
E_WhirlingB2_79.wav	Whirling

Table 7: The 58 referent sounds (part 1) used in [LR14]. All sounds were custom-made recordings.

Filename	Description
S_AM.wav	Basic synthesis example
S_AhOhAh.wav	Basic synthesis example
S_BandNoise.wav	Basic synthesis example
S_Click.wav	Basic synthesis example
S_Creee.wav	Basic synthesis example
S_Donwardring.wav	Basic synthesis example
S_HarshBubbles.wav	Basic synthesis example
S_LowIrregularPuff.wav	Basic synthesis example
S_ModulatedTriangle.wav	Basic synthesis example
S_RingNoise1.wav	Basic synthesis example
S_Sine.wav	Basic synthesis example
S_Square_Distorted.wav	Basic synthesis example
S_Sweep.wav	Basic synthesis example
S_UpwardScale.wav	Basic synthesis example
U_Accendigas_2.wav	Gaz stove lighter
U_BoccaAria#02.wav	A pen rubbed against a ventilation grid
U_Bottiglietta_Acqua#01.wav	A water bottle
U_Cappello#02.wav	A rubbed rubbed against a baseball cap
U_Cutter_1.wav	A cutter
U_Fiammiferi#01.wav	Matches being broken
U_Ombrello e strisce #01.wav	Some plastic pieces rubbed against an umbrella
U_Ombrello#02.wav	An umbrella
U_Polistirolo#01.wav	A piece of styrofoam
U_Porta_camera#02.wav	The door of the recording booth
U_Poster#01.wav	A poster roll
U_Poster#02.wav	A poster roll
U_Sedia#02.wav	A seat
U_Tubo_Plastica#03.wav	A poster tube

Table 8: The 58 referent sounds (part 2) used in [LR14]. All sounds were custom-made recordings or synthesis.

Folder	Subject	Gender
Subject10_IF/Edited	Subject 10	female
Subject11_IM/Edited	Subject 11	male
Subject12_IF/Edited	Subject 12	female
Subject14_IF/Edited	Subject 14	female
Subject16_IM/Edited	Subject 16	male
Subject18_IM/Edited	Subject 18	male
Subject3_IM/Edited	Subject 3	male
Subject7_IF/Edited	Subject 7	female

Table 9: The eight folder corresponding to the imitations of the eight subjects in [LR14].

6 Appendix: Contribution to Fonetik 2014

Sound initiation and source types in human imitations of sounds

Pétur Helgason
Department of Linguistics & Philology, Uppsala University, Sweden, and
Speech, Music and Hearing, KTH, Sweden
petur.helgason@lingfil.uu.se

Abstract

There exists a rich body of research exploring the production of speech, but for non-linguistic sound production, for example imitations of environmental sounds or animals, much less data and research are available. Data from human sound imitations collected in the initial, exploratory phase of the SkAT-VG project were analyzed in terms of the articulatory and aerodynamic conditions involved in their production. These exploratory data yielded a classification of sound productions in imitations based on the intersections between sound initiation and sound source types. The source types identified are turbulent, myoelastic, whistled and percussive sources. The ways in which these source types intersect with pulmonic, glottalic and velaric sound initiation, both egressive and ingressive, are described and discussed.

Introduction

In speech, the principal way of producing sound is to drive an airstream past one or more obstacles. The organ responsible for driving the airstream is the initiator (Pike 1943: 85ff), while the source of the sound produced is located at the point of the obstacle(s).

The sound initiation mechanisms commonly acknowledged in speech production are pulmonic egressive, glottalic egressive, glottalic ingressive and velaric ingressive (ibid.; see also Catford, 1977). Although there are no attested cases of pulmonic ingressive and velaric egressive airstreams being utilized as features in phonological systems, there is no real obstacle to producing sounds using these initiation mechanisms. Pulmonic ingressive sounds, in particular, are quite common (cf. Eklund, 2008), and also occur in imitations. Sounds can also be produced without creating an airstream, e.g. by clashing the teeth together or by slapping the tongue against the floor of the mouth. Such sounds are referred to as percussives (Pike, 1943: 103). Percussives are encountered in sound imitations, but they are rarely found in (non-pathological) speech.

The source-filter model of speech production (Fant, 1960) has been successful in describing the acoustics of human speech sound production. In speech the principal sound sources are voicing, produced with a pulmonic egressive airstream entraining the vocal folds into vibration, and friction noise, produced by constricting a pulmonic egressive airstream at some point in the vocal tract, causing turbulence. However, humans can produce sounds with a number of additional source types, some of which are used in spoken languages and some of which are not.

Sound initiation and source types

Here, the focus is on cataloguing source types that seem useful for sound imitation. The approach is to categorize the source types according to the articulatory and aerodynamic conditions under which they are produced. The main categories of source types thus identified are **myoelastic**, **turbulent**, **whistled** and **percussive**. The three former source types can be produced using various initiation mechanisms, but percussives constitute an initiation mechanism on their own. In the following,

examples of these four basic types of sources will be discussed primarily in terms of the initiation mechanisms involved and their observed or potential uses in sound imitation.

The exploratory data have various sources. Many of the examples on which the analyses are based have been found on-line, but exploratory recordings have also been made, with the aid of a professional improvisational actor.

Turbulent sources

To produce fricative sounds, an airstream is made turbulent by channeling it through a constriction in the glottis or the vocal tract (cf. Stevens 1999: 37f for an overview). In the exploratory phase of the SkAT-VG project we have observed imitations using pulmonic egressive friction (which parallel fricatives in speech) as well as velaric ingressive friction (which parallel clicks or click-like sounds). We have no examples yet where imitators use pulmonic ingressive, glottalic egressive or ingressive or velaric egressive friction.

Pulmonic egressive turbulence

Friction made with a pulmonic egressive airstream is by far the most commonly occurring turbulent source in imitations, just as it is in speech. As is the case with speech sounds, a turbulent friction noise can be made at many places in the vocal tract. This type of friction is especially common in the imitation of "basic" sound events, such as the interaction of solids (e.g. knocking, scraping and squeaking sounds) and sounds of gases in motion (e.g. blowing, puffing and hissing sounds) (cf. Lemaitre et al. 2011 for further examples of sound events). For example, the impression given by an improvisational actor of the sound of "scraping on a hard surface" is quite speech-like and can be described as a voiceless velar fricative [x].

Pulmonic ingressive turbulence

While pulmonic ingressive friction is not difficult to produce, it is difficult (or impossible) to produce sibilant fricatives with an ingressive airstream (Catford, 1988: 20ff; see also Eklund 2008 for a more comprehensive review). In other cases, although appreciably different, the acoustic result of ingressive friction is still quite similar acoustically to the egressive counterpart. These facts may contribute to its apparent scarcity in imitations. However, one should note that ingressive friction is encountered in emotive sounds, e.g. sucking in air through one's teeth to indicate pain (Cruttenden 1986: 180).

Glottalic egressive turbulence

Glottalic egressive friction is fairly common in languages, but as yet unattested in our exploratory data of imitations. Possibly, the acoustically similar outcomes of glottalic and pulmonic egressive friction are a contributing factor – why use a glottalic airstream when a pulmonic airstream creates, more or less, the same sound?

Glottalic ingressive turbulence

According to UPSID (Maddieson and Precoda, 1990) voiceless glottalic ingressive speech sounds (i.e., voiceless implosives) are phonologically distinctive in less than 1% of the world's languages. Judging by this typological rarity one could assume that such sounds are fairly difficult to produce. The exploratory data have not yet yielded imitations that make use of a glottalic ingressive airstream, as such. However, note Pike's (1943: 40) observation that English speakers sometimes use a voiceless velar implosive [ƙ] to imitate the "glug-glug" sound of pouring liquid from a bottle (the voiced counterpart can also be used). Thus, despite the typological rarity of such sounds, they still seem to be used in imitations.

Velaric egressive turbulence

A velaric egressive source has not been encountered in the exploratory data, but one can conceive of such sounds being used to imitate sputtering in liquids.

Squeezing a velaric airstream out between the teeth, for example, may faithfully replicate the sound of a spraying can (although, obviously, this depends on denture). An ingressive airstream leads to an acoustically similar result.

Velaric ingressive turbulence

Velaric ingressive turbulence is used to produce click sounds, which are typologically rare. Still, paralinguistic click sounds are encountered quite frequently in speech (cf., e.g., Jakobson, 1979: 40). In English, for example, the dental click even has a more or less standardized orthography, variably written as *tut-tut* or *tsk-tsk*.

In the SkAT-VG exploratory data set, the impression of "trickling water" made by an improvisational actor contains an example of velaric ingressive initiation (see Figure 1). To achieve this effect, the actor alternated soft postalveolar or alveolar click sounds with sublaminal percussives (discussed below, in the section on percussives) with frequent and rapid labial modifications of the resonance characteristics.

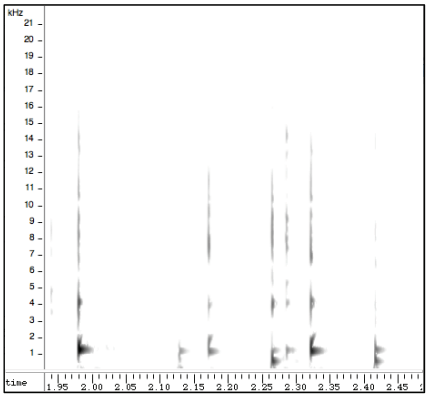

Figure 1. A spectrogram of an actor's impression of the sound of "trickling water".

Myoelastic sources

In the myoelastic source type, muscle and elastic tissue are made to oscillate in an air stream. This can lead to (almost) periodic sounds or intermittent breaks in an otherwise turbulent airstream. Crucially, for some myoelastic source types the oscillation is frequent enough to be perceived as a tone.

Pulmonic egressive myoelastic sources

The most commonly encountered myoelastic source by far, both in speech and sound imitations, is pulmonic egressive vocal fold phonation, i.e., voicing. As a sound source in speech and singing, the vocal folds are highly versatile, allowing a great deal of precision in the control of onset and offset, timbre and oscillation frequency.

In linguistic phonetics, a distinction is made between several vocal fold phonation types. Modal voice, breathy voice and creaky voice are the principal types (stiff voice, slack voice are also recognized but are not considered here, nor is the difference between breathy voice and whispery voice; see Ladefoged & Maddieson (1996) for an overview of the linguistic uses of voice). Non-linguistic voicing types include falsetto and pressed voice.

These various voice qualities are relevant for sound imitations, perhaps most notably in the imitation of animal sounds and engine sounds. The imitation of a cow, for example, usually involves a modal voice quality with a nasal resonance. The croak of a frog may be imitated with a creaky voice quality (an ingressive creak works even better). Falsetto voice is frequently encountered in animal imitations, e.g. when imitating a cat meowing.

A much less common myoelastic source type is aryepiglottic phonation in which the aryepiglottic folds vibrate in an air stream at frequencies ranging from approximately 40 to 100 Hz (Moisik, Esling & Crevier-Buchman, 2010). In the exploratory data we have observed impressions of animal growling in which aryepiglottic phonation is used, but usually it is used in combination with voicing. Similarly, there are examples of imitations of rumbling engines, which combine aryepiglottic vibration and voicing.

At least four types of supralaryngeal pulmonic egressive myoelastic sources

can be created. First, some people can achieve a uvular myoelastic oscillation, equivalent to uttering a voiceless, uvular trill, [ʀ̥]. Second, some people can achieve an apico-alveolar oscillation, equivalent to producing a voiceless, apico-alveolar trill, [r̥]. For these two source types, the rate of oscillation can exceed 30 Hz, but they are still not perceived as tones but rather as a rapid series of impacts. There are no examples of these two source types being used on their own for imitations in the exploratory data, but there are examples of the apico-alveolar source combined with whistling in bird imitations.

A third supralaryngeal source type uses a dorso-lateral configuration for the tongue and pushes out air between the tongue dorsum and a stricture that appears to be located at or anterior to the palatoglossal arch. The sound produced is periodic with an f0 range from approximately 150 and 700 Hz, judging from the examples gathered so far. The most well known use of the dorso-lateral source type is the voice of Donald Duck, the famous cartoon character. The exploratory data set contains numerous examples of the use of this source type in the imitation of birds.

The fourth supralaryngeal source type is made with a bilabial constriction. The constriction can be made with two distinct lip configurations, which yield quite different results. First, the lips can be pressed together without much stiffness in the labial tissue while an airstream is passed through. This leads to a fairly slow periodic myoelastic vibration (25-35 Hz) that is not perceived as a tone. The exploratory data set contains an example of such a voiceless, bilabial trill being used to imitate the blowing sound of a horse. The second lip configuration involves pressing the lips together quite tightly and making them much stiffer while forcing an airstream between them. This can lead to a (multiply) periodic source, which, in the exploratory data set, is found in the imitation of an elephant trumpeting.

Pulmonic ingressive myoelastic sources

When vocal fold phonation is made with a pulmonic ingressive airstream the result is ingressive voicing. Acoustically, ingressive voicing is quite distinct from egressive voicing, sounding harsher and less sonorant (cf. Eklund, 2008). Like egressive voicing, ingressive voicing can be made both as ingressive falsetto and ingressive creak.

In imitations, an ingressive falsetto is quite common. It is used to imitate various animal sounds, such as a dog bark, a pig squeal and crow caw, but it can also be used to imitate squeaking sounds, such as the squeaking sound of wiping a window pane.

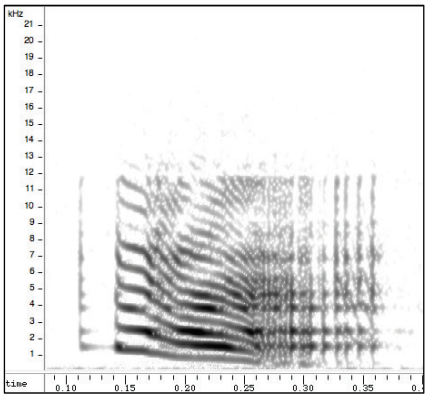

Figure 2. A spectrogram of an actor's impression of a "squeak from a window pane".

One example in the exploratory data set, shown in Figure 2, does contain both ingressive falsetto and ingressive creak. This is the impression made by an improvisational actor of the sound of a "squeak from a window pane".

Glottalic and velaric myoelastic sources

The SkAT-VG exploratory data set contains no imitations that make use of glottalic and velaric airstreams coupled with a myoelastic source. Using glottalic and velaric airstreams there is a very limited volume of air available to drive a myoelastic oscillation. Some configurations do yield a myoelastic effect, for example a glottalic egressive airstream can be coupled with an apico-alveolar

source to produce the equivalent of an ejective trill, [r̰']. However, the fact that these types of sources cannot be sustained for very long, if it can be achieved at all, reduces their usefulness in imitations, which may explain their absence in the exploratory data set.

Whistled sources

Very few languages are reported to have distinctive whistled coronal sibilants (Shosted, 2006). According to Shosted (ibid.: 566), whistled sibilants are produced in a manner similar to "a form of recreational whistling referred to as 'palatal' or 'roof' whistling", which is achieved by letting the tongue tip form a constriction that directs the airflow to the edges of the teeth. Pure "palatal" whistling is seldom encountered except in the repertoire of whistling virtuosi, such as the Hungarian Hacki Tamás or the Australian Luke Janssen. Still, the exploratory data set does include an example of this type of whistling being used to imitate the American Robin (*Turdus migratorius*).

In languages that do not have distinctive sibilant whistling, whistling can still occur sporadically when apical sibilants are produced and sibilants with a whistled component, similar to those found in speech, are observed when people imitate wind or weather noise.

Labial whistling does not occur in speech but the majority of people appear to be able to produce some form a labial whistle and this type of whistling is encountered frequently in daily life. Typically, labial whistling is pulmonic egressive, but it can almost as easily be produced ingressively. The exploratory data set contains examples of whistling being used to imitate birds, only in the form of palatal whistling and "digitally assisted" whistling (i.e. finger whistling), possibly because these generate higher oscillation frequencies.

Also, short labial whistling noises can be produced using both glottalic and velaric initiation, again both egressively and ingressively. The exploratory data contain several examples where imitators produce a short whistle with velaric egressive airstream to imitate the impact sound of a drop of water.

Percussive initiation

Percussive initiation does not require an airstream but results instead from an impact between solids, for example when the upper and lower teeth are made to clash or scrape together (Catford 1977: 63).

Percussives occur very rarely in (non-pathological) speech and are not phonologically distinctive in any language. Sands, Maddieson & Ladefoged (1993: 183) observe that, very rarely, an allophonic variant of an alveolar click is a percussive in which "the normal click is quite quiet but the tongue tip makes a forceful contact with the bottom of the mouth after the release of the front click closure". Incidentally, they also mention that this is a "sound sometimes made by speakers of non-click languages trying to imitate the sound made by the shoes of a trotting horse" (ibid.). As we saw in connection with Figure 1, the SkAT-VG exploratory data contain an example of such a "floored", sublaminal percussive, used as part of an impression given by an improvisational actor of "trickling water".

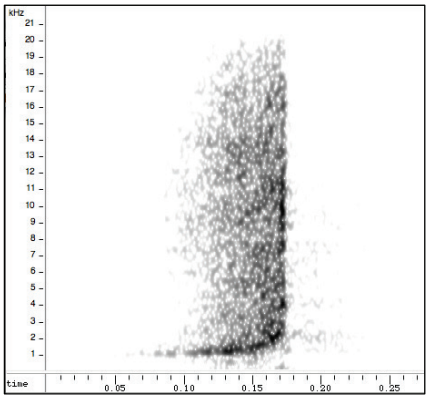

Figure 3. A spectrogram of an actor's impression of the sound of a "whip lash".

The data set also contains an example of a lamino-dental percussive, in which the tongue is shot forward at a

high velocity creating an impact sound as the lamina makes contact with the teeth and the alveolar ridge. This occurred in an improvisational actor's impression of the sound of a "whip lash", shown in the spectrogram in Figure 3. In speech, oral stop sounds are made at the offset of an occlusion by releasing a turbulent airstream through a narrow channel, giving rise to a high energy release burst. By contrast, in the example in Figure 3, the "burst" at 0.17 ms in the spectrogram is created at the onset of the occlusion and is in fact the sound of the impact of the tongue lamina against the teeth.

Conclusion

The observations made during the exploratory phase of the SkAT-VG project have shown that in sound imitations humans can utilize a far wider range of articulations than are used to make phonological distinctions in languages. Also, imitators can utilize sound initiation mechanisms and source types that are not part of the repertoire of their native language(s) and in many cases they utilize mechanisms that are typologically rare (and considered "difficult").

A classification of sound productions is proposed that is based on three basic source types, turbulent, myoelastic and whistled, intersecting with six basic sound initiation mechanisms, pulmonic, glottalic and velaric initiation, both egressive and ingressive. In addition, percussive sounds form a class of their own, being both an initiation mechanism and a source type.

Acknowledgements

I wish to thank Prof. Sten Ternström for helpful advice and comments. This research has been supported by EU-FET grant SkAT-VG 618067.

References

Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Catford, J. C. (1988). *A practical introduction to phonetics*. Oxford: Clarendon Press.

Eklund, R. (2008). "Pulmonic ingressive phonation: diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech." *Journal of the International Phonetic Association 38*: 235-324.

Fant G. (1960). *The Acoustic Theory of Speech Production*. The Hague: Moulton.

Jakobson, R. and Waugh, L. (1979). *The Sound Shape of Language*. Bloomington, Ind: Indiana University Press; and London: Harvester.

Lemaitre, G., Dessein, A., Susini, P. & Aura, K. (2011). "Vocal Imitations and the Identification of Sound Events". *Ecological Psychology, 23*:4, 267-307.

Maddieson, I. & Precoda K. (1990) Updating UPSID. *UCLA Working Papers in Phonetics 74*, pp. 104-111.

Moisik, S. R., Esling, J. H., & Crevier-Buchman, L. (2010). A high-speed laryngoscopic investigation of aryepiglottic trilling. *Journal of the Acoustical Society of America 127 (3)*. 1548-1559.

Pike, Kenneth L. (1943). *Phonetics: a critical analysis of phonetic theory and a technic for the practical description of sounds*. Ann Arbor, MI: University of Michigan Press.

Sands, B., Maddieson, I. and Ladefoged, P. (1993). "The Phonetic Structures of Hadza". *UCLA Working Papers in Phonetics, No. 84: Fieldwork Studies in Targeted Languages*.

Shosted, Ryan. 2006. "Just put your lips together and blow? Whistled fricatives in Southern Bantu." In H. C. Yehia, D. Demolin & R. Laboissiere (Eds.), *Proceedings of ISSP 2006,* (pp. 565–572). Belo Horizonte: CEFALA.

Stevens, K. N. (1999). *Acoustic Phonetics*. Cambridge: MIT Press.

References

[Hel14] Pétur Helgason. Sound initiation and source types in human imitations of sounds. In *Proc. Fonetik*, 2014. Submitted.

[LDSA11] Guillaume Lemaitre, A. Dessein, Patrick Susini, and K. Aura. Vocal imitations and the identification of sound events. *Ecological psychology*, 23(4):267–307, 2011.

[LR14] Guillaume Lemaitre and Davide Rocchesso. Vocal imitations communicate sounds more effectively than verbalizations. *Journal of the Acoustical Society of America*, 135(2):862–873, 2014.

[New04] Fred Newman. *MouthSounds: How to Whistle, Pop, Boing, and Honk... for All Occasions and Then Some*. Workman Publishing, 2004.