Perspectives on soundscapes

Verb constructions in descriptions of everyday sounds

Eleni Tzimopoulou, Jenny Hartman, Joost van de Weijer, and Carita Paradis
Lund University

This study explores how listeners use verb constructions to describe everyday sounds. It considers (i) how the verb constructions provide Gestalt to events in the form of activities, processes and states, and (ii) how the constructions correspond to different aspects of a soundscape (sound sources, acoustic properties, and listener reactions). We found that verb constructions primarily describe sound sources and that these are portrayed as dynamic events. Acoustic properties were rarely described through verb constructions. When they were, the constructions primarily conveyed static properties of sound as in *the sound is loud* as opposed to dynamic events as in *the sound comes closer*. Listener reactions were primarily expressed through comparative constructions such as *it sounds like* and justifications such as *I can hear*. The upshot of the study is that speakers conceptualise sounds according to highly dynamic Gestalts for which the main focus is the source of the sound.

Keywords: acousmatic sounds, auditory experiences, conceptual structures, event constructions, sound descriptions

1. Introduction

When people experience everyday sounds, they associate their acoustic properties (e.g., loudness, acoustic texture) with source events and entities that can be found in the sounds' environments (e.g., Brown et al., 2016; Guastavino, 2021; Nudds, 2009). The term *soundscape* (Southworth, 1969) is sometimes used to refer to people's conceptualizations of sounds as experiences of the physical world (e.g., Kang & Schulte-Fortkamp, 2016; Schulte-Fortkamp et al., 2023). Hartman and Paradis (2023) introduced an approach to researching soundscapes through the

https://doi.org/10.1075/rcl.00233.tzi | Published online: 22 September 2025
Review of Cognitive Linguistics ISSN 1877-9751 | E-ISSN 1877-976X

lens of language. They asked participants to freely describe 20 *acousmatic* stimuli, i.e., sounds heard without visual support (e.g., Dhomont, 1996; Kane, 2014). These stimuli included everyday sounds drawn from different environments such as sounds generated by human activity (e.g., eating and walking) and sounds of nature (e.g., rain and bird song). Hartman and Paradis (2023) focused on how the participants' descriptions evoke meanings within the conceptual domains of EVENT (source of the sound), SOUND (acoustic properties), and EXPERIENCE (listener reaction), and on how these descriptions reflect the participants' perspectives on the sounds as elements of the soundscape. They showed that accounts of a source event are central to the descriptions of sound.

The present study builds on these findings. The aim is to establish how sound experiences are described through verb constructions and how these constructions are distributed across different sounds. It also aims to establish how perceivers of sounds construe the EVENTS in terms of ACTIVITIES, PROCESSES and STATES in relation to the different types of sounds and to the three elements of soundscapes, i.e. the source event, the sound itself and the perceiver. Note that this use of the term event refers to event structures in language, not to be confused with source event as an element of the soundscape by Hartman and Paradis (2023).

The article is structured as follows. Section 2 introduces previous work on sound, both as a phenomenon and an experience, outlines the findings of previous studies that have focused on representations of sounds in discourse, and provides a brief overview of the cognitive semantic views on conceptual structures and the construal of perspective. The methodology of the study is presented in Section 3. Section 4 presents the results, and Section 5 discusses the findings. Section 6 concludes the article.

2. Background

Section 2.1 describes how the terms *sound, soundscape*, and *acousmatic experience* are used in this article. Section 2.2 summarizes previous research about sound descriptions, specifically descriptions of sound-related EVENTS in fiction narratives and written descriptions of acousmatic stimuli. Section 2.3 focuses on how perspectives are made salient in communication.

2.1 Sounds, soundscapes, and acousmatic experiences

Everyday situations can give rise to different types of sounds, and it has been proposed in the literature that these sounds often translate into different types of lis-

tener experiences. Distinctions have been made, for instance, between *musical sounds* such as the ringing of an alarm clock (Gaver & Norman, 1988) and *environmental sounds* such as someone making breakfast (e.g., Dubois et al., 2021; Marcell et al., 2007). When listeners experience musical sounds, their attention is often geared towards the *acoustic properties, physical properties*, or *musical parameters* of the sounds such as pitch, loudness, or duration (e.g., Dubois, 2000; Gaver & Norman, 1988). When listeners experience environmental sounds, on the other hand, they instead attend to what is happening, what is causing the sounds, and where the sounds are coming from (e.g., Gygi et al., 2007; Houix et al., 2012; Jenkins, 1985; Marcell et al., 2000; Pasnau, 1999; Rosenblum et al., 1996).

Everyday sounds are frequently generated by human activities, such as people talking, crying, laughing, or moving in different ways. Sounds generated by human activity make up an inclusive category referred to as *anthrophonic sounds* (e.g., Kang & Schulte-Fortkamp, 2016). Anthrophonic sounds include background noise, such as distant incomprehensible chatter (e.g., Samodra, 2018) as well as sounds generated by the human body without involvement of the vocal cords, such as a munch or a footstep (e.g., Brown et al., 2016; Brown et al., 2011; Cox, 2008). Research has shown that anthrophonic sounds are meaningful constituents of everyday experiences (e.g., Cox, 2008; Hartman & Paradis, 2023).

An *acoustic environment* is essentially the collection of all sounds within a setting at a specific place and time (e.g., Fernström & Taylor, 2019; Pijanowski et al., 2011). A *soundscape*, on the other hand, while also potentially associated with a combination of coexisting sound sources (e.g., Bergman et al., 2009; Bizley & Cohen, 2013; Bregman, 1994; Marcell et al., 2007), stems from experiences of acoustic environments. The notion of a soundscape is therefore an experiential/conceptual construct. The subsequent excerpt summarizes how soundscapes have been defined, as cited in Fiebig (2023, p. 27).

> The international standard on soundscape (ISO 12913-1) defines soundscape as a perceptual construct that must be distinguished from the physical phenomenon: "Soundscape is an acoustic environment as perceived or experienced by a person or people in context." (ISO 12913-1 2014, p. 1).

The distinction between acoustic environments and soundscapes as experiences of these environments stems from the fundamental distinction between "the nature" of sounds as phenomena in the physical world and "the content" of auditory experiences (O'Callaghan & Nudds, 2009). Soundscapes originate in a variety of physical environments (Brown et al., 2011; Guastavino, 2021). Urban environments, for instance, have been associated with auditory experiences of sounds stemming from both human activity and non-human events (Brown et al., 2011). Human activity in this case includes anthrophonic and mechanical sounds

from traffic or uses of machines and devices. Non-human events may involve domesticated animals, wildlife, natural elements such as water or other substances and materials, and phenomena like thunder (Brown et al., 2011; Guastavino, 2021).

People who have experienced an acoustic environment are later able to re-create it via *simulation* and integrate different aspects of their experience into one conceptual unit (e.g., Barsalou et al., 2004; Gibbs Jr & Matlock, 2008; Hartman & Paradis, 2018). Situated simulations of everyday acoustic environments are often influenced by the type of connections to the physical world the listeners derived from a sound at the moment of listening (e.g., Hartman & Paradis, 2023; Steenson & Rodger, 2015; Tuuri & Eerola, 2012) and may focus on different elements of a soundscape. Figure 1 illustrates a soundscape that stems from the sound of digging. It includes three elements: the source (*event*) that generates the sound, the acoustic properties that are generated (*sound*), and the *listener* who is engaged as an active perceiver. The illustration of a soundscape in Figure 1 aligns with common approaches in research on soundscapes, which consider the listener not merely as a passive receiver of sounds, but as an active participant within "a dynamic system of information exchange" (Fiebig, 2023, p. 23).

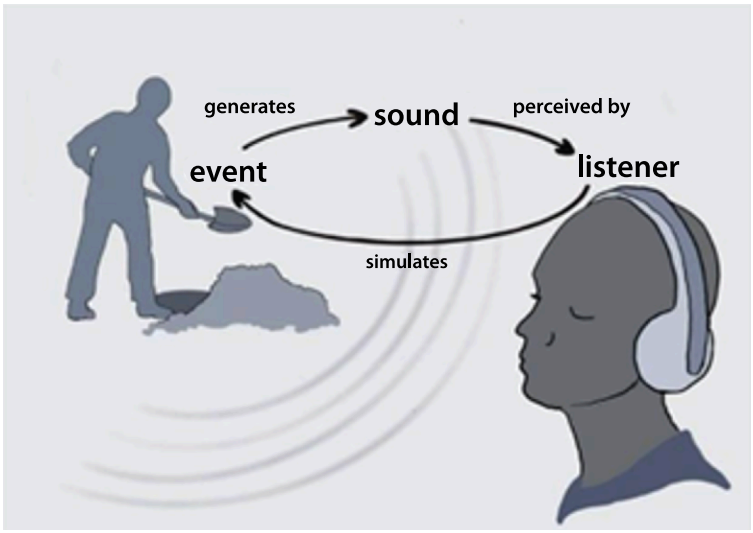

Figure 1. Elements of the soundscape of digging

In the study by Hartman and Paradis (2023), participants described soundscapes such as the one shown in Figure 1 by focusing on a single element (e.g., *someone digging gravel with a shovel*), on two elements (e.g., *Someone digging with a metal shovel. There is crunching from the earth and sometimes resonance*

from the spade.), or on all three elements of the soundscape (e.g., *A person using a shovel to dig up earth or possibly snow after a snowfall. I can hear the shovel hitting the ground and lifting.)*.

Notably, soundscapes are variable and contextually dependent. Even familiar sounds can be experienced differently by a single listener, depending on the situation, expectations, or purposes that drive attention to the sounds. Schulte-Fortkamp and Fiebig (2023) emphasize this point by stating that "if humans are perceiving and considering their acoustic environments, they are also cognizant of the overall "gestalt" of all sensory components of that environment" (p. 5). Acoustic properties influence listeners' responses to sounds, but these responses can also be affected by several facets of everyday life such as the perceiver's gender, sensory acuities or sensitivities, social etiquette, cultural contexts, or survival instincts (Cox, 2008). For instance, psychoacoustic studies have shown that scraping sounds such as a tool scratching a hard surface (Halpern et al., 1986) and anthrophonic sounds such as crying, coughing, snoring, and munching are often perceived as highly unpleasant (Cox, 2008). Sounds such as water movements, bird-chirping, and background chatter, on the other hand, can have a positive effect on the listener.

2.2 Representations in discourse of sounds as EVENTS

Descriptions of sound have been investigated from a lexical perspective in studies focusing, for example, on iconicity (e.g., Akita & Parudeshi, 2019; Dingemanse, 2012; Winter et al., 2017) and sensory meanings associated with all the senses (e.g., Caballero et al., 2019; Winter, 2019). These studies address sensory meanings evoked by adjectives and nouns, while research on how verbs describe sound is scarce and deals primarily with verbs as stand-alone lexical items. For instance, in his work on how verbs reflect sensory meanings, Viberg (2019) distinguished between "experiencer-based" verbs such as *listen to* and *hear*, and "phenomenon-based" verbs such as *crack* and *rattle* (p. 19).

Only two previous studies have focused on how sounds are described in discourse, namely Caballero and Paradis (2020) and Hartman and Paradis (2023). These two studies are summarized in the following.

Caballero and Paradis (2020) investigated how sounds are represented in fiction narratives. They found that sounds are often conceptualized as EVENTS whose descriptions are instantiated in domains other than sound proper, as in *strong men banging their way in and out of the café* (Caballero & Paradis, 2020, p. 723). The example shows how different aspects of EVENT meanings (such as AGENT, ACTION, MOTION and SPACE) combine with meanings of SOUND in discourse. Caballero and Paradis (2020, p. 716) showed that when sounds are

described as EVENTS, these events are contextualised either from the point of view of sound as a phenomenon (as in *the sound boomed like an immense bell stroke*), the point of view of the producer of the sound (as in *some peasant woman banging on the gate*), or the experiencer's viewpoint (as in *she heard the beat of the drums*). Of these three points of view, descriptions of sound as a phenomenon were the least frequent. The study highlights the importance of context for a proper understanding of how sounds are represented in discourse (Caballero & Paradis, 2020, p. 707).

Research that taps into participants' acousmatic experiences has verified that people often approach sounds contextually. Hartman and Paradis (2023) investigated participant descriptions by focusing on meanings associated with different perspectives on sound: distal perspectives associated with EVENT meanings such as AGENT and MANNER, medial perspectives associated with meanings of SOUND such as QUALITY and TEMPORAL STRUCTURE, and proximal perspectives associated with EXPERIENCE meanings such as FAMILIARITY and EMOTION (p. 9). The results showed that most of the descriptions contained EVENT meanings. Meanings of SOUND were primarily used in combination with EVENT descriptions, as sound being in motion (e.g., *jumpy sound* (p. 12)) or action (e.g., *the sound bites* (p. 17)). When descriptions reflected a single perspective, this perspective was usually distal. Meanings of EXPERIENCE were the least frequent across the descriptions. In other words, the participants predominantly focused on the sound sources (p. 12).

The findings of these two studies align with previous research showing that "environmental sounds are dynamic" and "convey action and movement-related information" (Marcell et al., 2000, p. 833). In the current study, we investigate in more detail how conceptualizations of EVENTS are expressed in sound descriptions in the form of verb constructions. We also investigate how verb constructions reflect distinct perspectives on sound.

2.3 Conceptual structures, construals, and constructions

In order to identify the dynamic meanings that linguistic structures evoke in communication, it is essential to distinguish between conceptual structures (ontologies) and cognitive processes (construals) as shown in Table 1 (Paradis, 2005, 2015). Ontologies, on the one hand, are structures of content proper such as concrete entities or activities, and configurations are structuring templates such as boundaries of forces. Construals, on the other hand, operate on the conceptual structures to establish discursive meaning that is relevant to a particular situational context. Examples of construals that shape communication adopt a particular perspective and focus on specific aspects of an experience are perspective,

Gestalt formation and focus of attention to make something salient (e.g., Croft & Cruse, 2004; Paradis, 2005). In language use, a speaker adopts a specific viewpoint (*perspective*), a focus point (*salience*) and a profiling (*Gestalt*) of an experience, and thus activates combinations of conceptual structures through construals that constrain meaning by focusing on the aspects of a given experience that are significant to a given situation.

Table 1. Ontologies and construals (adapted from Paradis, 2015, p. 278)

Ontologies (conceptual structures)		**Construals (cognitive processes)**
Content	**Configurations**	
concrete phenomena	boundary	perspective
action, process, state	thing, relation	salience
abstract phenomena	deixis	Gestalt
	scale	
	force	
	path	
	figure/ground	

Finally, meanings are expressed in communication through linguistic constructions including the syntactic, lexical, semantic, and pragmatic properties of the words or phrases as a set of features that create a message in a particular context (Fillmore, 2006; Fillmore et al., 1988; Zadrozny et al., 1995). *Constructions* are "form-meaning correspondences" (Goldberg, 1995, p. 1), where *meaning* "is to be construed broadly enough to include contexts of use" (p. 229).

2.4 The present study

In the present study, we focus on the construal of perspective to capture listeners' descriptions of soundscape elements. We believe that the verb constructions used in the descriptions reflect structures of the sounds' source events, as well as the participants' perceptions of and reactions to the sounds' acoustic properties. We focus on the meanings of verb constructions because they reflect how language users conceptualize soundscape elements as EVENTS. We investigate the frequencies of the constructions and also include the type of Gestalt profiling they evoke in terms of ACTIVITIES, PROCESSES, and STATES. The acousmatic sounds that we investigate stem from human and non-human sources. For these two sound types, we compare the perspectives listeners adopt when they describe different elements of the corresponding soundscapes, as well as how these elements are structured as EVENTS in the descriptions. Our consideration of previous research on

perceptions of acoustic environments as soundscapes (reviewed in Section 2) has guided us in asking the following questions.

1. Are there differences in the frequencies with which the participants use verb constructions to describe the sounds. If so, can these differences be correlated to the sounds described? For instance, if verb construction frequency correlates with the degree of perceived activity, we may expect human-caused sounds to elicit more verb constructions than non-human sounds.
2. How often do participants use verb constructions to describe the sounds in terms of dynamic events (processes or activities) or static events (states)? Can frequencies of verb constructions expressing dynamic and static events be correlated to the sounds described?
3. What are the specific types of verb constructions (e.g., [X DO Y]) that express dynamic events, and what are the specific types of verb constructions that express states (e.g., [X BE Y])?
4. How often and in what form are verb constructions used to provide salience to the sources of the sounds, the sounds themselves, and the listener's experience across the sound descriptions?

3. Methods

The following sections describe the data used for the study (Section 3.1), the annotation of the descriptions (Section 3.2), and the data analysis (Section 3.3).

3.1 Materials

We used textual data compiled by Hartman and Paradis (2023), available at the Open Science Framework (https://osf.io/xv52g/). In this study, 214 participants provided descriptions of 20 everyday, acousmatic sounds downloaded from storyblocks.com (https://www.storyblocks.com/audio/sound-effects). The participants were recruited via social media and Prolific (https://www.prolific.com/). They were at least 18 years old and had English as their native language. Importantly, the participants were not provided with any information about the sounds. They were asked to respond to an open-ended prompt (i.e., "Describe the sound in as much detail as possible, as if you are describing it to someone who cannot hear it") which did not direct their attention towards specific aspects of the sounds, allowing them to describe each sound freely. The task was self-paced and completed by the participants online. The texts written by the participants are not edited in any way, which means that there are many sentence fragments and

anacolutha, and hence the data share many traits with everyday conversation data or informal writing (Põldvere et al., 2021) The sounds were presented in random order, followed by a textbox in which the participants typed their description.

The sounds are listed in Table 2. The left column in the table provides labels used in this article. The middle column offers brief descriptions of the sounds retrieved from the database, slightly altered to include human activities, non-human events, objects, materials, and substances that are audible in the audio-clips. The right column indicates the source events.

Table 2. Overview of sounds

Label	Description	Type of source events
snore	continuous snoring	sounds featuring vocalizations
cough	continuous coughing	
baby	a baby crying/laughing	
crowd	a crowd cheering and clapping	
apple	munching an apple	sounds originating in the mouth
chip	crunching chips	
tea	sipping tea and exhaling	
teeth	brushing teeth	
onion	chopping vegetables	sounds stemming from objects, materials, and substances manipulated by the hands
textile	tearing fabric	
dishes	washing dishes	
water	moving in water	
digging	digging and shovelling earth materials	
gravel	walking on gravel	sounds stemming from movements of the feet on different types of surfaces
heels	walking on stairs wearing heels, in enclosed space with audible echo	
forest	sounds of nature from birds	ambient sounds stemming from non-human events
rain	rain and thunder	
fire	a fireplace burning	
traffic	sounds of downtown traffic	ambient sounds featuring mechanical noises
grocery	sounds in a grocery store from the store counter, incomprehensible chatter, objects such as shopping bags	

3.2 Identification of verb constructions

The first step of the data annotation was to identify all constructions headed by verbs. The verbs were used as hubs for the whole constructions in the practical

processing of the data. In this section, we give examples of straightforward cases of verb constructions as well as less straightforward instances that are always part and parcel of data of this kind. In the annotation phase, the verbs were annotated using their base forms as shown in (1).[1]

(1) *Someone vigorously* <u>*brushing*</u>[brush] *teeth,* <u>*sounds like*</u>[sound] <u>*scrubbing*</u>[scrub] *but* <u>*moving*</u>[move] *around and echoey as it's*[be] *inside a mouth.* <u>*Sounds*</u>[sound] *wet* (/teeth/:60)

Since the texts were written in an informal style much like everyday spoken language, there are cases that out of context could be understood in more than one way. Clear instances were annotated by one member of the group, while less straightforward cases were discussed and resolved by at least two members of the group. In (1), *scrubbing* is such a case. We interpreted it as an anacoluthon that serves as a reinforcement of *brushing* and that is structurally paralell with *someone brushing teeth* and *moving*, i.e., 'sounds like someone is scrubbing, hence a verb construction.

Constructions including light verbs such as *make, do*, or *have* were annotated in the base form of the verb as in (2) and (3). The meanings evoked by the longer constructions headed by such verbs were subequently interpreted in the contexts where they were used.

(2) *[…] in a tub* <u>*having*</u> [have] *a bath […]* (/water/:9)

(3) *[…] the knife* <u>*makes*</u> [make] *some resistance […]* (/onion/:119)

EVENT constructions which occurred in isolation as in (4) were identified as verb constructions, but instances of EVENT meanings construed as nouns or adjectival descriptors modifying nouns, as the items appearing in bold in Examples (5) and (6), were not included.

(4) <u>*Coughing*</u>[cough] (/cough/:35)

(5) ***Pouring raining*** *and* ***rumbling*** *thunder. Constant* ***raining****, very heavy and very wet with possible* ***flooding****. Occasional thunder* (/rain/:21)

(6) *The* ***breaking up*** *of dry top-soil with a sharp spade or shovel.* (/digging/:14)

1. Annotated elements in the examples are underlined, followed by their annotations given in superscript. The parenthetical labels at the end of each example contain the sound and the participant number.

3.3 Verb coding scheme

To link EVENT structures and perspectives, we designed a coding scheme based on the Lexical Meaning as Ontologies and Construals (LOC) framework (Paradis, 2005). As described in Section 2.3, this framework enables a distinction between conceptual structures and construal operations, which is essential when we investigate how speakers create meaning and evoke perspectives on sounds in line with their experiences.

Table 3. Annotation scheme

EVENT types	EVENT structures by verb constructions	Perspective
ACTIVITIES	transitive activity, active [X DO Y]	**source** (source event)
PROCESSES	intransitive activity, active [X ACT]	
	passive [X UNDERGO]	
	inherently intransitive, active [X HAPPEN]	**sound** (acoustic properties)
STATES	source match [X BE Y]	
	reasoning [X BE LIKE Y]	
	reaction [X EXPERIENCE Y]	**listener** (reactions)
	possession, feature attribution [X HAVE Y]	

An overview of the coding scheme is shown in Table 3. The left column of this table shows three EVENT types, namely ACTIVITIES, PROCESSES and STATES. The first two of these represent ongoing events and were treated as a single category in our analysis. They include a scale from agentive expression with a clear configuration of force from the X to Y to low force happenings. The third type was used for static events. Examples (7a) and (8a) contain a variety of instances of both types.

(7a) *This is*[STATE] *the sound of somebody walking*[ACTIVITY/PROCESS] *on gravel, there is*[STATE] *a slight echo as they walk*[ACTIVITY/PROCESS] *over the stones that crunch*[ACTIVITY/PROCESS] *under their feet. They are walking*[ACTIVITY/PROCESS] *quite fast but with even paces.* (/gravel/:160)

(8a) *The sound of gravel underneath someone's feet. Crickets can be heard*[STATE] *in the distance. As well as the hoots of owls. There's*[STATE] *another sound that I can't make out*[STATE] *as well, which sounds like*[STATE] *the footsteps of a dog? or horse?* (/gravel/:168)

The EVENT types were then analysed and categorized according to the labels shown in the middle column of Table 3. ACTIVITIES/PROCESSES includes transitive [X DO Y], intransitive [X ACT], passive structures [X UNDERGO] or [X HAPPEN]. STATES were categorized as matches between sounds and their sources [X BE Y],

structures evoking reasoning [X BE LIKE Y], constructions headed by sensory, emotional or cognitive meanings (e.g., *hear, feel, hate, assume, guess, imagine*) [X EXPERIENCE Y], and constructions that express possession or [X HAVE Y]. To illustrate the categories, the Examples (7b) and (8b) show how they were applied to the annotated verb constructions in Examples (7a) and (8a).

(7b) [**X** (the trigger) **BE** (is) **Y** (the sound of somebody)] [**X** (somebody) **ACT**((is) walking)] [**X** (a slight echo) **BE** (is there)]
[**X** (they) **ACT** (walk)] [**X HAPPEN** (the stones crunch)] [**X** (they) **ACT** (are walking)]

(8b) [**X** (the listener) **EXPERIENCE** (can hear) **Y** (the crickets)] [**X** (another sound) **BE** (is there)]
[**X** (the listener) **EXPERIENCE** (can't make out) **Y** (another sound)] [**X** (another sound) **BE LIKE** (sounds like) **Y** (the footsteps of a dog or horse)]

The right column of Table 3, finally, shows three perspectives on sound (i.e., source, sound, listener). These perspectives were categorized as mutually exclusive and included constructions describing source events (focus on *source*), acoustic properties (focus on *sound*), or listeners' reactions (focus on *listener*). Examples (9) to (11) are descriptions of the sound of fire written by three different participants.

(9) *… fire burning something …* (/fire/:6)

(10) *… splinters of sound … often coming in ones or twos* (/fire/:140)

(11) *… I feel warm* (/fire/:39)

While the descriptions are of the same sound, the first one focuses on the source of the sound, the second one on acoustic properties of the sound, and the third one on the experience that the sound elicits in the participant.

3.4 Data analysis

To answer the questions presented in Section 2.4, we present and describe the distributions of the various categories in the same order in which they were presented above. Regression models were used to estimate the frequencies of the categories. The frequencies of the verb constructions were estimated using Poisson regression, in which we control for the variation in the number of words that participants used to describe the set of sounds. Binomial regression models were subsequently used to estimate proportions of the verb construction categories listed in Table 3. All analyses were performed using the software R (version 4.0.2; R Core Team, 2020), with the additional package lme4 (Bates et al., 2015). The graphs that show the results of the analyses were made with the package ggplot2

(Wickham, 2016). In Section 4, we first provide estimates of verb construction frequencies across the 20 sounds. We then estimate how many of these constructions expressed ACTIVITY/PROCESS and how many expressed STATE. Subsequently, we provide the distribution of the four ACTIVITY/PROCESS types and the four STATE types, presented in Table 3. Finally, we present the distribution of the three perspectives.

4. Results

The data included 3,875 descriptions consisting of 51,089 words in total. Table 4 gives an overview. The average description length varied roughly from 8 to 20 words, and the relatively high numbers of the longest descriptions suggest that the distribution of the numbers of words was positively skewed.

Table 4. Material overview

Sound	Description length (words)	
	Average	Range
grocery	20.4	2–86
heels	16.5	2–87
forest	16.4	1–78
traffic	16.3	1–238
gravel	15.6	2–261
digging	15.3	1–243
dishes	15.0	1–58
onion	14.0	1–75
rain	13.4	1–54
water	12.8	1–48
fire	12.2	1–72
tea	11.4	1–60
crowd	11.2	2–61
chip	10.8	1–51
teeth	10.6	2–59
apple	10.4	1–63
baby	10.4	2–68
textile	9.7	1–41
snore	8.8	1–39
cough	8.4	1–120

Section 4.1 begins with an overview of the verb construction frequencies across the sounds. Section 4.2 covers EVENT types; it details the proportions of verb constructions that realize ACTIVITIES/PROCESSES and STATES across the sounds. Sections 4.3 and 4.4 detail the proportions of the eight groups of constructions realizing different EVENT structures (see Table 3). Section 4.5 details the proportions of verb constructions that realize three distinct perspectives on the sounds, namely the focus on sound source(s), the sound itself, and the listener and ends with a discussion of the verb construction types in relation to the perspectives

4.1 Frequencies of verb constructions

There were 467 descriptions with no instances of constructions headed by verbs. In the remaining 3,813 descriptions, 8,374 verb constructions were found. The frequencies of the verb constructions across the sounds ranged from 313 in descriptions of /rain/ to 637 in descriptions of /dishes/. Figure 2 illustrates the estimated verb construction frequencies for every 10 words on the horizontal axis across the 20 sounds on the vertical axis, sorted in descending order based on the estimates. For rain, at the bottom of the graph, a little bit more than one in every 10 words is expected to be a verb. In contrast, for baby, at the top of the graph, this rate is nearly twice as high, or one in every five words is expected to be a verb.

The legend of Figure 2 indicates shared features of the sounds' sources as shown in Table 2. The estimates at the upper end of the scale consisted of vocal sounds and sounds originating in the mouth (except /teeth/). Those at the lower end of the scale were mechanical sounds, non-human sounds and sounds associated with feet movements. Interestingly, in between the upper and the lower ends were mostly sounds produced by hands.

4.2 Event types

Out of all verb constructions in the dataset, 5,609 (67%) described an ACTIVITY/PROCESS, while the remaining 2,765 (33%) described a STATE. Figure 3 illustrates the estimated proportions of ACTIVITY/PROCESS constructions for each sound. As in Figure 2, the sounds on the vertical axis have been sorted in descending order based on the estimates.

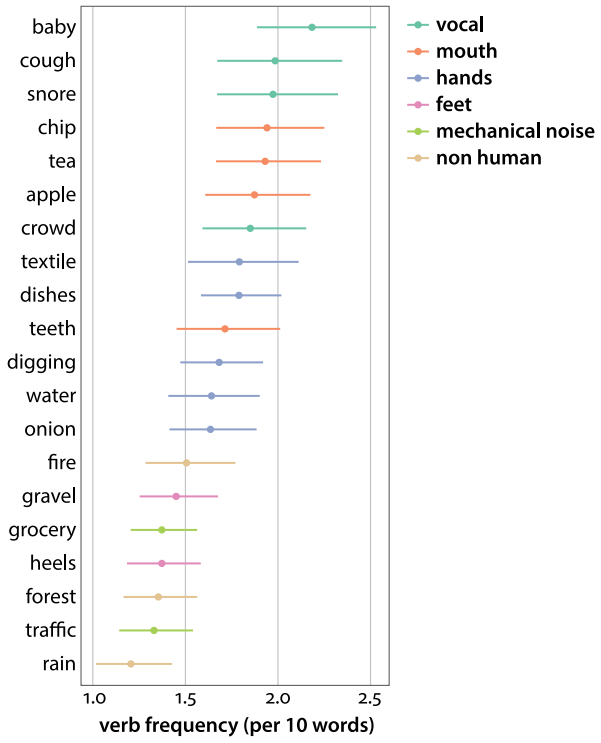

Figure 2. Estimated verb construction rates per 10 words
Error bars are 95% confidence intervals

Figure 3 shows that the estimated proportions of ACTIVITY/PROCESS constructions exceed one half of all verb constructions. As in Figure 2, the estimates of ACTIVITY/PROCESS constructions that describe sounds from non-human events and feet movements are lower than those from human sources. The ordering of the sounds, however, is similar but not the same. Sounds that are at the top of Figure 2 have moved further down in Figure 3. Notably, vocal sounds are not as frequently described as an ACTIVITY/PROCESS, as for instance, sounds produced by the mouth.

4.3 ACTIVITY/PROCESS constructions

As indicated in Table 3, the ACTIVITY/PROCESS constructions were divided into four types ([X ACT], [X DO Y], [X HAPPEN] and [X UNDERGO]). Figure 4 shows how these four types were distributed across the 20 sounds. The sounds are listed alphabetically on the vertical axis.

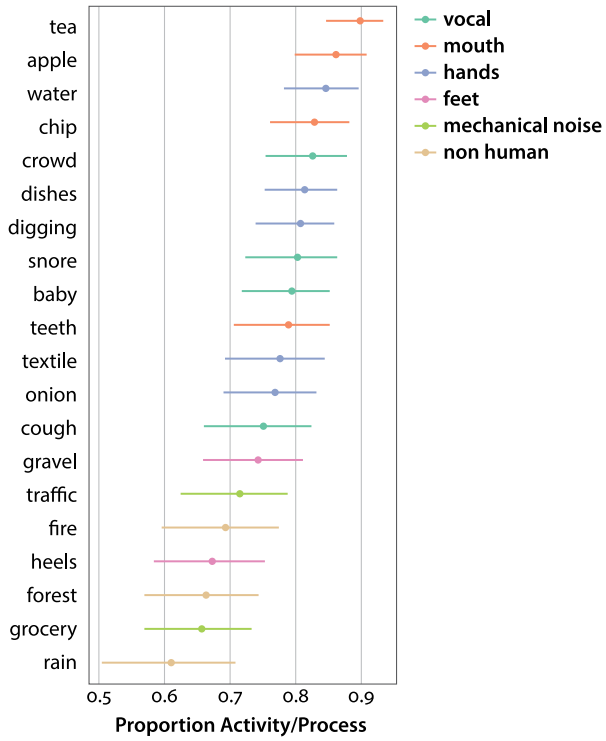

Figure 3. Estimated proportions Activity/Process constructions
Error bars are 95% confidence intervals

Figure 4 shows that for many sounds, one of the four types (mostly [x ACT] and [x HAPPEN]) clearly dominates the remaining three categories. Vocal sounds and sounds associated with movement of feet were primarily described as intransitive activities [x ACT]. Sounds originating in the mouth and sounds associated with hand movements were primarily described as transitive activities [x DO y]. Sounds stemming from non-human events displayed high proportions of different structures: /rain/, /fire/, and /traffic/ were primarily described as processes [x HAPPEN]; /forest/ and /grocery/ were primarily described as intransitive activities [x ACT]. Passive structures expressed as [x UNDERGO] constructions had higher proportions for /dishes/, /grocery/, and /textile/ compared to all other sounds, for most of which the proportion of [x UNDERGO] was equal to or very close to zero.

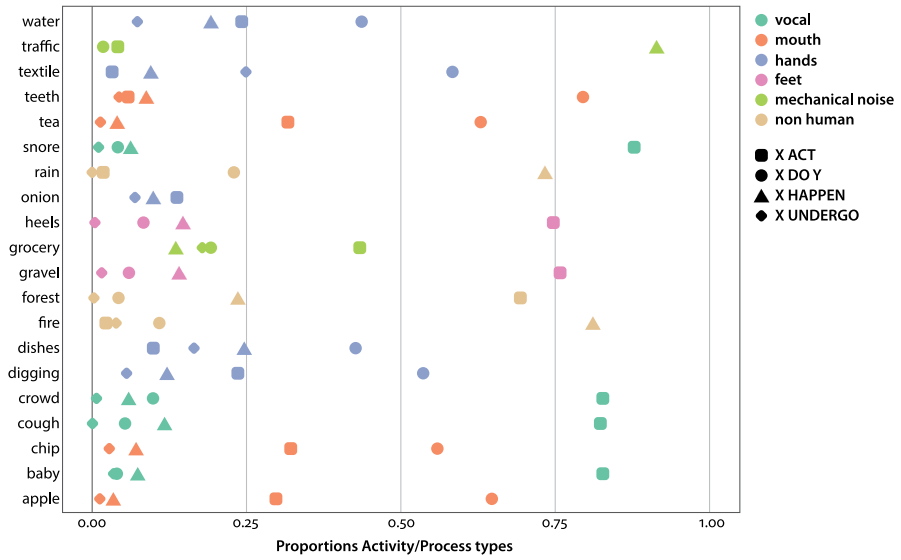

Figure 4. Distribution of types of activity/process constructions

4.4 State constructions

Figure 5 illustrates the estimated proportions of [x be y], [x be like y], [x experience y], and [x have y] of the total frequency of state constructions for each sound.

Unlike the four types of activity/process constructions, there were no clearly dominant types of state constructions for any of the sounds. Nearly all proportions were between 0.0 and 0.5. In other words, the categories within state constructions were more equally distributed than the categories within activity/process constructions. The proportion of the [x be y] constructions was close to 0.50 for most sounds, except for /water/, /textile/, /dishes/, /chip/, and /apple/. For these sounds, state was primarily realized as [x be like y]. There is an almost equal distribution of the [x experience y] constructions across the sounds, with proportions close to 0.25. The [x have y] construction was mainly described for /heels/, /cough/, and /gravel/.

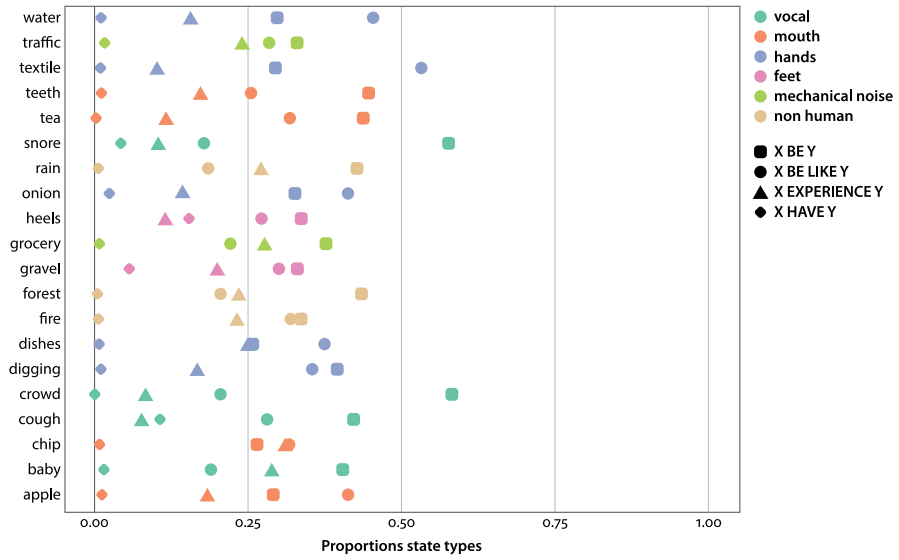

Figure 5. Distribution of state construction types

4.5 Perspectives

The three perspectives on sound that we investigated established a focus on the source, the sound itself (acoustic properties), or the effect of the sound on the listener (cf. Table 3). Figure 6 illustrates the estimated proportions of these three perspectives across the sounds.

As Figure 6 shows, the estimated proportions of verb constructions that describe sources were high (more than 0.75) across all sounds, and the participants rarely focused on acoustic properties of the sounds. In other words, participants focussed primarily on the sources of the sounds. They sometimes, but not often, mentioned effects that the sounds had on them (i.e., their subjective experiences), but they hardly ever mentioned the sounds' acoustic properties.

Table 3 lists the eight types of verb construction that occur in the data. The types of verb construction used to describe *source* were primarily construed as ACTIVITY/PROCESS and only a few STATES: 32% [X ACT], 29% [X DO Y], 17% [X HAPPEN], 15% [X BE Y], 6% [X UNDERGO], and 1% [X HAVE Y].

The construction types describing *listener* reactions were primarily STATES: 55% [X BE LIKE Y] constructions, 41% [X EXPERIENCE Y], and 2.6% [X BE Y]. The [X UNDERGO], [X HAPPEN], [X DO Y], [X ACT], and [X HAVE Y] constructions had percentages lower than 1% each. The [X EXPERIENCE Y] structures were expressed by constructions evoking meanings of thought, justification, and emotional reaction, as in *what I would assume to be a piece of fruit* (/apple/:22), *you can tell by the*

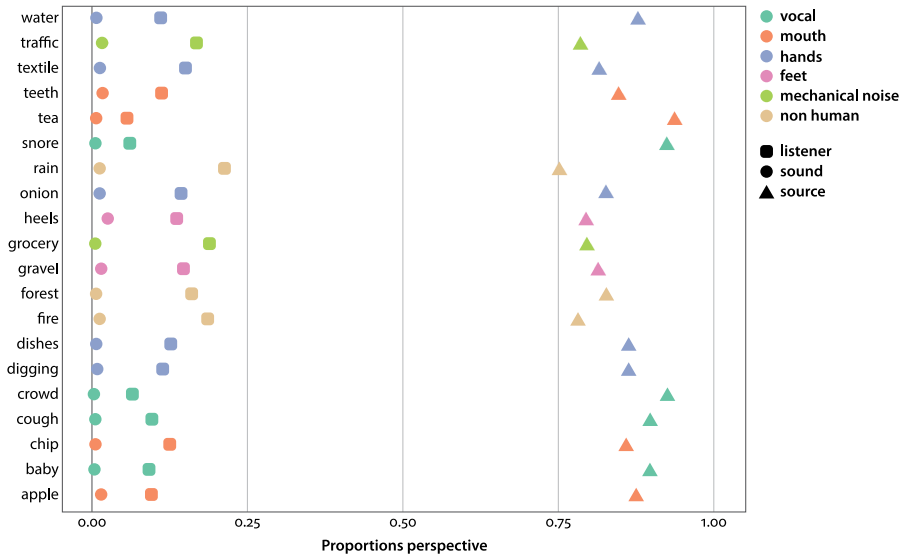

Figure 6. Distribution of perspectives taken by the participants

way the feet... (/heels/:206), and *I would love to be in this sound* (/forest/:124). Justifications such as *I can hear the clunking of the shovel* (/digging/:1) and *you can hear the rain* (/rain/:98) dominated the descriptions expressed as [X EXPERIENCE Y]. The frequencies of these constructions were higher in descriptions of sounds stemming from non-human events such as /rain/ and /forest/, and lower in descriptions of vocal sounds such as /snore/ and /baby/. Verb constructions that evoke emotional reactions have a low frequency across the data and are only sporadically used in the descriptions of 12 out of 20 sounds.

Regarding the types of verb constructions that describe the *sound* itself across the data, we found that 63% were construed as STATES [X BE Y] and 31% expressed as PROCESSES [X HAPPEN] across all sounds. The complement of the verb in [X BE Y] were either adjectives or nouns as in *the sound overall is quite loud* (/storm/:17) and as in *there's a crunch at the beginning* (/apple/:155). Examples of [X HAPPEN] are *the sound rapidly gets louder* (traffic/:95) and *a boom unfolding into a bass roll* (/storm/:14). Sporadically, the [X HAVE Y] constructions, as in *the sound has a deep, basey hum* (/forest/:50), and the [X UNDERGO] constructions were used, as in *a brief sound followed by a thud* (/onion/:50).

5. Discussion

The specific focus in this study has been on the different types of verb constructions used by participants to describe acousmatic sounds to people who cannot hear them. It has focused on the distribution of verb constructions across a range of sounds and the types of Gestalt profiling of content meanings as ACTIVITIES/PROCESSES and STATES, and construction types such as [X ACT], [X DO Y], and [X BE LIKE Y]. The analysis also included the construals of three different perspectives that may be made salient in the soundscapes, i.e., the source, the sound and the listener reaction. The findings are important in that they demonstrate that experiences of sounds are primarily described as dynamic events rather than as states. In a sense this is a natural consequence of the fact that an occurrence of a sound involves a sound emitter at one end of a spatio-temporal path to a perceiver at the other.

Previous research has shown that when people experience everyday sounds such as those examined in this study, they try to identify the events causing the sounds in specific environments and contexts and has also considered how sounds are interpreted and categorized by listeners (e.g., Gygi et al., 2007; Brown et al., 2011; Jenkins, 1985; Marcell et al., 2000). To the best of our knowledge, our study is the first one to investigate how language users attribute EVENT structures to sounds through verb constructions. Our analysis showed that sounds are predominantly construed as ACTIVITIES/PROCESSES across sounds from distinctly different sources such as vocal sounds, mechanical sounds, and sounds from nature. These findings are in line with previous research indicating that sound descriptions are dynamic (e.g., Caballero & Paradis, 2020; Hartman & Paradis, 2023; Marcell et al., 2000; Winter 2019). We also observe that descriptions of sounds stemming from similar source events have similar proportions of ACTIVITY/PROCESS constructions. We found that the proportions of four distinct groups of verb constructions that evoke meanings of ACTIVITY/PROCESS structures in the descriptions are particularly common in descriptions of the source of the sounds. In the case of vocal sounds and sounds stemming from movements of different body parts, those EVENT structures highlight the force of the vocalizations and body movements. In the case of sounds stemming from non-human events, the descriptions feature both transitive and intransitive ACTIVITY/PROCESS EVENT structures, as in *fire breaking down the wood* (/fire/:106), *birds chirping cheerfully* (/forest/:1), *rainfall splattering against the ground* (/rain/:1) and *cars passing from left to right* (/traffic/:1).

The analysis of STATE constructions showed that these constructions are not very often used to describe the sources of the sounds. Only 16% of the STATE constructions were found to describe the sources as in *there is also a lawnmower in*

the background (/digging/:11) and *this is in a supermarket* (/grocery/:134). Across the descriptions, STATE constructions primarily express the participants' reasoning about the sounds and less often express emotional responses. The low proportions of verb constructions that focus on the sounds as phenomena in our study indicate that the sounds themselves are not often accounted for in the form of verb constructions. When they are, however, the structures that describe them are mainly of two types, [X BE (Y)] and [X HAPPEN]. The proportions are higher for [X BE (Y)] indicating that verb constructions match sounds to their attributes (nominalization of EVENT) more often than they describe sounds as ACTIVITY/PROCESS structures.

6. Conclusion

This analysis is based on over 8,000 verb constructions in approximately 3,500 descriptions of 20 everyday, acousmatic sounds by 214 participants. The aim of the study was to find out what verb constructions were used to describe the sounds, and how they were distributed across the different sounds and soundscape elements. The study paid special attention to how the event structures expressed by the verb constructions were construed in terms of ACTIVITY/PROCESS and STATE Gestalts and how they were distributed across the sounds and the elements of the soundscape. On the basis of the distinct event structures that the constructions described in context, they were subdivided into eight construction types: [X ACT], [X DO Y], [X HAPPEN], and [X UNDERGO] for ACTIVITY/PROCESS constructions; [X BE (Y)], [X BE LIKE Y], [X EXPERIENCE Y], and [X HAVE Y] for STATE constructions.

The results offer new insights into this previously unexplored topic, showing that when participants describe everyday sounds to people who cannot hear them, they do not use verb constructions at the same rate across different sounds. Sounds stemming from human activity are described using verb constructions more often than sounds stemming from non-human events. The specific body parts that are associated with the source events also play a role in the verb construction frequency in the descriptions. Specifically, verb construction frequency in descriptions of vocal sounds and sounds associated with mouth and hands is higher compared to descriptions of sounds associated with movements of feet, whose descriptions contain some of the lowest verb construction frequencies in the dataset.

The results also offer new knowledge about the Gestalt profiling of events. The analysis showed most of the constructions are construed as activities and

processes rather than states. This dynamic aspect of soundscapes is prominent regardless of whether the sounds stem from human activity or non-human events.

When verb constructions describe sounds stemming from different sources, they evoke distinct event structures, primarily reflecting the structures of the source events specific to each sound. When static events are described, the verb constructions primarily express introspective meanings, most often through comparisons as in *it sounds like x* and justifications as in *I can hear the x* and least often through emotional reactions as in *I feel x* or *I (dis)like x*. This use of constructions in the form of comparisons indicates that.

In sum, we know very little about how perception of sound is represented in language, and therefore work on how humans perceive sounds and describe them through language is crucial for a better understanding of this relationship. In this study, we put the spotlight on this issue, showing that people primarily describe the sources of sounds as dynamic events. We argue that this is due to the fact that sounds are phenomenologically inherently dynamic with a sound emitter and sound waves that hit the perceiver's ears. Our findings contribute to the theoretical discussion of how sounds are conceptualized as events, detailing the event structures evoked through language in terms of their Gestalts, and identifying different types of verb constructions that participants attribute to soundscape elements.

Funding

This research was funded by the Kamprad Family Foundation (grant number 20190045). We would also like to thank two anonymous reviewers for constructive feedback and Sara Farshchi for reading and commenting on the manuscript at an earlier stage.
Open Access publication of this article was funded through a Transformative Agreement with Lund University.

References

Akita, K., & Parudeshi, P. (2019). *Ideophones, mimetics and expressives*. Amsterdam: John Benjamins.

Barsalou, L., Sloman, S., & Chaigneau, S. (2004). The HIPE theory of function. In L. Carlson & E. van der Zee (Eds.), *Functional features in language and space: Insights from perception, categorization, and development* (pp. 131–148). Oxford University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bergman, P., Skold, A., Vastfjall, D., & Fransson, N. (2009). Perceptual and emotional categorization of sound. *Journal of the Acoustical Society of America*, *126*(6), 3156–3167.

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, *14*(10), 693–707.

Bregman, A.S. (1994). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT press.

Brown, A.L., Gjestland, T., & Dubois, D. (2016). Acoustic environments and soundscapes. In J. Kang & B. Schulte-Fortkamp (Eds.), *Soundscape and the built environment* (pp. 1–16). Boca Raton: Taylor & Francis Group, LLC.

Brown, A.L., Kang, J., & Gjestland, T. (2011). Towards standardization in soundscape preference assessment. *Applied Acoustics*, *72*(6), 387–392.

Caballero, R., & Paradis, C. (2020). Soundscapes in English and Spanish: A corpus investigation of verb constructions. *Language and Cognition*, *12*(4), 705–728.

Caballero, R., Paradis, C., & Suárez-Toste, E. (2019). *Representing wine: sensory perceptions, communication and cultures*. Amsterdam: John Benjamins.

Cox, T.J. (2008). Scraping sounds and disgusting noises. *Applied Acoustics*, *69*(12), 1195–1204.

Croft, W., & Cruse, D.A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

Dhomont, F. (1996). Is there a Québec sound? *Organised Sound*, *1*(1), 23–28.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistic Compass*, *6*(10), 654–672.

Dubois, D. (2000). Categories as acts of meaning: The case of categories in olfaction and audition. *Cognitive Science Quarterly*, *1*(1), 35–68.

Dubois, D., Cance, C., Coler, M., Paté, A., & Guastavino, C. (2021). *Sensory Experiences: Exploring Meaning and the Senses*. Amsterdam: John Benjamins.

Fernström, M., & Taylor, S. (2019). An uncomfortable audio ethnography: Sound and politics in the evolution of an acouscenic listening approach to Softday's Sonically Engaged Art Practice. *Acoustic Ecology Review*, *18*(1).

Fiebig, A. (2023). Soundscape: A construct of human perception. In B. Schulte-Fortkamp, A. Fiebig, J.A. Sisneros, A.N. Popper & R.R. Fay (Eds.), *Soundscapes: Humans and their acoustic environment* (pp. 23–48). Cham: Springer.

Fillmore, C.J. (2006). Frame semantics. In D. Geeraerts (Ed.), *Cognitive linguistics: Basic readings* (pp. 373–400). Berlin: Mouton de Gruyter. (Reprinted from Linguistics in the Morning Calm, Linguistic Society of Korea (Ed.), (pp. 111–137). Seoul: Hanshin Publishing Company).

Fillmore, C.J., Kay, P., & O'Connor, M.C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, *64*(3), 501–538.

Gaver, W.W., & Norman, D.A. (1988). *Everyday listening and auditory icons*. University of California, San Diego, Department of Cognitive Science and Psychology.

Gibbs Jr, R.W., & Matlock, T. (2008). Metaphor, imagination, and simulation: Psycholinguistic evidence. In R.W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 161–176). Cambridge: Cambridge University Press.

Goldberg, A.E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Guastavino, C. (2021). Exploring soundscapes. In D. Dubois, C. Cance, M. Coler, A. Paté, & C. Guastavino (Eds.), *Sensory experiences: Exploring meaning and the senses* (pp. 139–167). Amsterdam: John Benjamins.

Gygi, B., Kidd, G.R., & Watson, C.S. (2007). Similarity and categorization of environmental sounds. *Perception & Psychophysics*, *69*(6), 839–855.

Halpern, D.L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception & Psychophysics*, *39*, 77–80.

Hartman, J., & Paradis, C. (2018). Emotive and sensory simulation through comparative construal. *Metaphor and Symbol*, *33*(2), 123–143.

Hartman, J., & Paradis, C. (2023). The language of sound: events and meaning multitasking of words. *Cognitive Linguistics*, *34*(3–4), 445–477.

Houix, O., Lemaitre, G., Misdariis, N., Susini, P., & Urdapilleta, I. (2012). A lexical analysis of environmental sound categories. *Journal of Experimental Psychology: Applied*, *18*(1), 52–80.

Jenkins, J.J. (1985). Acoustic information for objects, places, and events. *Persistence and change*, 115–138.

Kane, B. (2014). *Sound unseen: Acousmatic sound in theory and practice*. Oxford: Oxford University Press.

Kang, J., & Schulte-Fortkamp, B. (Eds.). (2016). *Soundscape and the built environment* (Vol. *525*). Boca Raton: Taylor & Francis Group, LLC.

Marcell, M., Malatanos, M., Leahy, C., & Comeaux, C. (2007). Identifying, rating, and remembering environmental sound events. *Behavior Research Methods*, *39*(3), 561–569.

Marcell, M.M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of clinical and experimental neuropsychology*, *22*(6), 830–864.

Nudds, M. (2009). Sounds and Space. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 0). Oxford: Oxford University Press.

O'Callaghan, C., & Nudds, M. (2009). Introduction: The philosophy of sounds and auditory perception. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays*. Oxford: Oxford University Press.

Paradis, C. (2005). Ontologies and construals in lexical semantics. *Axiomathes*, *15*, 541–573.

Paradis, C. (2015). Meanings of words: Theory and application. In H. Ulrike & S. Petra (Eds.), *Handbuch Wort und Wortschatz* (pp. 274–294). Berlin: De Gruyter.

Pasnau, R. (1999). What is sound. *The Philosophical Quarterly*, *49*, 309–324.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H., & Pieretti, N. (2011). Soundscape ecology: The science of sound in the landscape. *BioScience*, *61*(3), 203–216.

Põldvere, N., Johansson, V., & Paradis, C. (2021). On the London–Lund Corpus 2: Design, challenges and innovations. *English Language and Linguistics*, *25*(3), 459–83.

Rosenblum, L.D., Wuestefeld, A.P., & Anderson, K.L. (1996). Auditory reachability: An affordance approach to the perception of sound source distance. *Ecological psychology*, *8*(1), 1.

Samodra, F.T.B. (2018, 11–12 October 2017). Soundscape elaboration from anthrophonic adaptation of community noise. IOP Conference Series: Earth and Environmental Science, Medan, Indonysia.

Schulte-Fortkamp, B., Fiebig, A., Sisneros, J.A., N. Popper, A., & Fay, R.R. (Eds.). (2023). *Soundscapes: Humans and their acoustic environment*. Springer.

Southworth, M. (1969). The sonic environment of cities. *Environment and Behavior*, *1*(1), 49–70.

Steenson, C. J., & Rodger, M. W. (2015). Bringing sounds into use: Thinking of sounds as materials and a sketch of auditory affordances. In Q. U. B. School of Psychology (Ed.), (pp. 1–33).

Tuuri, K., & Eerola, T. (2012). Formulating a revised taxonomy for modes of listening. *Journal of new music research*, *41*(2), 137–152.

Viberg, Å. (2019). Phenomenon-Based Perception Verbs in Swedish from a Typological and Contrastive Perspective. *Syntaxe et Sémantique*, *20*(1), 17–48.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New Yord: Springer Verlag.

Winter, B. (2019). *Sensory linguistics: Language, perception and metaphor*. John Benjamins Publishing Company.

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? Iconicity in English sensory words. *Interaction Studies*, *18*(3), 443–464.

Zadrozny, W., Szummer, M., Jarecki, S., Johnson, D. E., & Morgenstern, L. (1995). NL Understanding with a Grammar of Constructions. In M. Nagao, International Conference on Computational Linguistics.

Address for correspondence

Carita Paradis
Centre for Languages and Literature
Lund University
Box 201
221 00 Lund
Sweden
carita.paradis@englund.lu.se

Biographical notes

Eleni Tzimopoulou is a PhD student of English Language and Linguistics at Lund University, Sweden. Her project concerns meanings of sensory language in natural language use. She has also worked as a research assistant for the *Language for sensory perceptions* project funded by the Kamprad Family Foundation and run by Jenny Hartman and Carita Paradis. https://www.sol.lu.se/en/person/EleniTzimopoulou/

https://orcid.org/0000-0001-8866-2460

Jenny Hartman is Associate Professor of English Language and Linguistics at Lund University, Sweden. Her research explores how diverse ways of experiencing and interacting with the world are reflected in language and communication. She takes a particular interest in how sensory experiences are conveyed, especially by individuals with Autism, Obsessive-Compulsive Disorder, and Tourette Syndrome, examining how they use language, gestures, and pictures to articulate their lived experiences. Jenny Hartman | Centre for Languages and Literature

https://orcid.org/0000-0001-6800-9050

Joost van de Weijer is a researcher at Lund University Humanities Lab in Sweden. He collaborates with various researchers on the topics of animal communication, prosody and factivity, and motivation in high school students to learn a foreign language. He teaches courses in psycholinguistics and in the statistical analysis of experimental data. https://www.humlab.lu.se/person/JoostvandeWeijer/

https://orcid.org/0000-0001-9843-3143

Carita Paradis is Professor of English Language and Linguistics at Lund University, Sweden. She has published extensively on the dynamics of meaning-making in human communication to contribute to a better understanding of what language reveals about social interaction, perception and cognition, and inversely how those systems influence and structure natural language use. https://www.sol.lu.se/en/person/CaritaParadis/

https://orcid.org/0000-0002-7240-9003

Publication history

Date received: 10 October 2023
Date accepted: 14 July 2025
Published online: 22 September 2025