1

How to Study and Develop Communication Acoustics

The rest of this book describes facts, theories, and models of communication by sound and voice – both human and machine. The objective of basic research in this field is to better understand how we communicate, while the engineering goal is to develop and use technology to make this communication more versatile and powerful. In both cases, the challenge is to understand a variety of topics and to solve problems by approaching them from several points of view.

1.1 Domains of Knowledge

Different phenomena in nature, living organisms, society, and technology obey different laws and exhibit different properties and thus require different scientific concepts to study them. Even when phenomena appear similar, a detailed interpretation of specific laws to be applied may be quite different. That is why we can have (and actually must use) different *domains of knowledge*.

In some problems it is enough to have a look at just one aspect of reality (one domain), but with complicated multidisciplinary problems, many domains must be mastered. In complex cases, such as the topic of this book, we are forced to know at least the basics of several domains of knowledge in order to master the field properly. Some of these domains are more basic, while some are important from a methodological point of view in research and others from an application or a practical point of view. For a modern approach to sound and voice communications, we deal in this book with the following domains of knowledge:

- *Acoustics* and the fundamentals of physical sound deal with the interaction between physical objects that form the basis for sound and vibration phenomena. This domain answers questions on how sound is generated by sound sources, how it propagates or is absorbed, how it behaves in a closed space such as a room, auditorium, or a concert hall, and so on.
- *Signal processing* comprises the theories and techniques on how signals carrying information are generated, transmitted, analysed, or transformed by humans, in nature, or using technical

Communication Acoustics: An Introduction to Speech, Audio, and Psychoacoustics, First Edition.
Ville Pulkki and Matti Karjalainen.

devices. In this domain, the basic physical details are of less importance, and abstraction from physical interaction to (typically) one-directional input–output causal relations is emphasized.
- *Speech and audio* are concerned with the fundamental characteristics of spoken and audio signals, and the specific techniques for communication with them. *Spoken language* and *speech or language technology* are of great importance, although in this book we will discuss them only briefly. *Music*, *music acoustics*, and *music technology* also belong to this category.
- *Physiology* and the *psychophysics (psychoacoustics) of hearing* study how our auditory system works both from a physiological and a functional point of view. This domain includes several important subtopics, one of which is *spatial hearing*, which deals with how we localize sound sources and perceive sound environments.
- *Mathematics* and *computer science* are the sciences that provide the general methodologies required to get the formal understanding necessary to model and realize complex communication processes.
- *Engineering applications* is the field that utilizes a wide variety of techniques for sound and voice communications. This includes traditional acoustics design, control of noise, speech technology (speech analysis, synthesis, coding, recognition), audio technology (recording, production, reproduction), multimedia sound (integration of sound with other modalities) virtual acoustics, and technical audiology for helping with hearing problems, etc.

Research in these and related topics is often scattered and separated. One of the main objectives of this book is to provide the reader with a broad overview, without digging too deeply into details of any subtopic, since there are more specialized books and publications available for that purpose. Before starting this journey of facts, theories, and models, a characterization of related research methodologies and how they evolve is given.

1.2 Methodology of Research and Development

Scientific and engineering knowledge of communication processes has developed over the last hundreds, even thousands, of years. First there were beliefs that sounded rational, but a closer investigation found them to be false or only partially true. Some early thinkers and experimenters, such as Pythagoras, turned out to be very successful and influential (Hunt, 1992). Pythagoras studied the behavior of strings and understood the inverse relation between pitch and string length in ancient Greece during the sixth century B.C. Related to the harmony (consonance) of sounds, he assumed that the whole universe followed simple number rules of harmony. In the 1700s experimental studies in, and mathematical theories of, the physical sciences improved our understanding on how, for example, sound propagates in the air. Sound sources were better understood, and Hermann von Helmholtz had a modern view on the functioning of the human ear (von Helmholtz, 1954) more than a century ago. Since then, electronic and digital communications have forced scientists and engineers to learn more deeply the secrets of signals and the information conveyed by them.

There are three basic ways a scientist or engineer may acquire knowledge about a system or process, such as communication by sound and voice:

- *Experimentation* with, say, concrete physical or psychophysical, systems. This has been the standard approach used by scientists and engineers, augmented by practical or theoretical

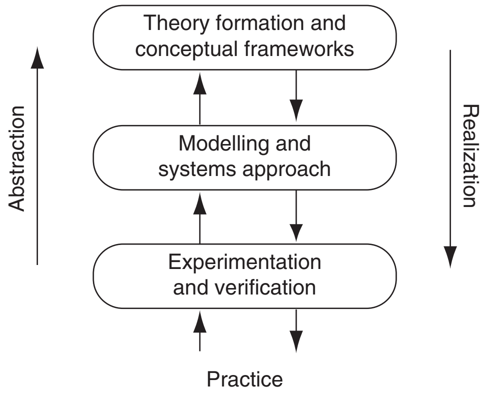

Figure 1.1 Abstraction and realization processes in research through experimentation, modelling, and theory formation.

thinking. Concrete experimentation involves a large number of details, which makes seeing general laws behind the multitude of phenomena difficult, but on the other hand it helps to verify that the thinking is not too far from reality.

- *Modelling* of an object or phenomenon by means of constructing another kind of system that simulates the original phenomenon or object but is easier to experiment with. The first modelling techniques were based on constructing concrete, physical objects as models. The modern approach, using computer-based simulation, has expanded the capability of modelling techniques dramatically.
- *Theory formation* is a conceptual approach to formulating general principles inherent in a specific field of research. A theory has higher generality and potentially the widest applicability compared to experimentation and modelling techniques. On the other hand, very general theories are hard to apply to complex and specific cases in practice; they rather help and direct the thinking process.

Many forms of scientific and engineering methodologies can be formulated as different combinations of the three approaches. Figure 1.1 characterizes these as levels of the abstraction process as well as realizing abstractions back towards practice. *Abstraction* means getting rid of details that are not essential for modelling or understanding a specific object or principle, and concretization (realization) is the opposite process of adding details in order to approach reality or practical construction of a system.

Recently, in this age of computers being applied everywhere, the three approaches of experimentation, modelling, and theory formation have started to merge and integrate. Experiments are often carried out using computational models before the final validation against practice. The difference between theory and computational models is also becoming less distinctive, since advanced computer models, including artificial intelligence and logic programming, may enable high-level theoretical inference.

Progress in research and practical applications is based on continuous iteration between the levels of abstraction, aiming at improved models and theories that better meet the challenges of reality.

1.3 Systems Approach to Modelling

A *systems approach* to studying a specific object means abstraction and modelling which is based on the finding that very many (if not all) objects in reality exhibit some general properties that can be described using *systemic concepts*. The target may be of a physical origin, a society of living beings, or a human-made technical artefact, and yet some general concepts may apply to all of them.

Figure 1.2 shows a very simple diagram and Figure 1.3 a more complex one that may be understood as systemic descriptions of something, not specific to any domain of knowledge until so related. Based on such descriptions, one may compile a list of concepts representing systemic aspects. The reader may wish to propose another set of systemic concepts, or disagree with the authors about the usefulness of such concepts. The list presented here is not intended as a basis of a 'general systems theory', rather it orients the reader to the way the authors have tried to formulate the organization and content of this book.

- *Element*: an entity that is considered not to have an internal systemic representation. This means that (a) we have an abstract element, for example in mathematics, (b) we don't want to pay attention to an element's internal complexity, or (c) we are not able to represent the element in more detail due to its internal complexity. A typical element representing an entity without an internal description but with an input–output relationship is the *black-box* diagram of Figure 1.2.
- *Relation*: something that represents a connection or link, like an interaction, similarity, or distance, between elements, objects, or systems. Types of relations will be mentioned for different systemic concepts below. A *property* can be understood as a relation of an object to

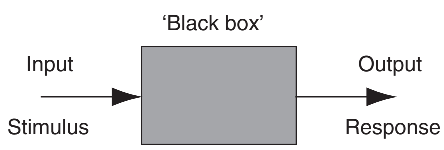

Figure 1.2 A black-box element with an input–output relationship.

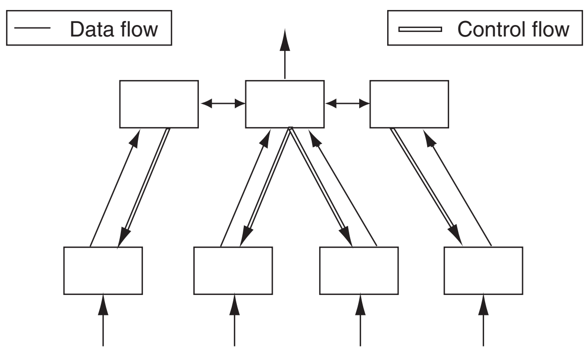

Figure 1.3 A block diagram characterizing a system that consists of elements, data, and control flow links between them, and a hierarchical-type (two-level) organization.

a general reference, such as a classification system. Causal relations are discussed with the concept of function(ality) below.

- *Structure*: represents a set of objects and relations between them that constitute a relatively permanent aspect of the system under study. The structure may define a spatial, temporally ordered, or a more abstract composition of entities and relations. Structures are often hierarchical (see below) through *parts* and *part-of* relations, i.e., an entity consists of parts that together form the (whole) entity.
- *Function(ality)*: an aspect that is complementary to structure, i.e., it represents relations between changing properties of the entity under study. *Causal relations*, meaning relations between cause and effect, stimulus and response, or input and output, are particularly important since the functionality of physically realizable systems is based on them, requiring that the (causal) effect cannot appear before the cause in time.
- *Event*: an activity in a system that occurs within a specific time span. A *stream* is a sequence of events.
- *State*: a concept that describes how a system ‘memorizes’ the history of its behaviour. Its reaction (output) to external causes (inputs) depends also on its internal state (unless it is a memoryless system).
- *Object*: a generic concept that has been used in many ways, for example as a target of observation or manipulation. From computer science the concept has gained a more systematic usage – object-oriented methodology – where an object is understood as an entity having an internal state and functionality through type-dependent functions.
- *Type* and *class*: concepts that are based on the similarity or dissimilarity of entities. Entities that share a common property or properties are categorized in the same type or class from the viewpoint of this property. In object-oriented methodology, a class is a definition for the structure and behaviour of its *instances*, i.e., objects for computation based on the class definition.
- *System*: a very general and generic concept to describe any object that has enough systemic properties, although in specific system theories it may be given a precise meaning. It typically refers to an abstract representation where the specific domain of knowledge, such as acoustics, is deliberately left aside and systemic concepts are used instead.
- *Control*: goal-driven behaviour. Living species have evolved and learned functionalities with target behaviour towards more successful life in their environment. Advanced machines are programmed or they adapt to their working conditions in order to meet some *optimal* criteria. *Feedback* is an important principle of control where the response is compared with the goal in order to improve control.
- *Process*: refers to (often controlled or goal-driven) the functionality of a system. Processes and processing are typically described as input–output relationships or data- and control-flow diagrams.
- *Organization*: represents a higher-level structure and functionality, the result of a goal-driven process. *Self-organization* refers to the ability of a system to improve its organization towards more advanced functionality. Real self-organization is more demanding and complex than simple *learning* and adaptation.
- *Hierarchy*: the *parts* and *part-of* relations in structures that constitute systems common in nature. A hierarchy in an organization means a level-like control structure where ‘lower-level’ units realize detailed activities and ‘higher-level’ units perform more strategic control, the flow of control taking place ‘vertically’ between levels. Figure 1.3 depicts a strictly hierarchical two-level system where control flows from top to bottom and data from the bottom

up. A *heterarchical* system is an organization principle where control is distributed and functional objects act in several organizations, representing different levels in different sub-organizations. A strict hierarchy may be very efficient in optimal conditions, but a heterarchy is more versatile, especially if its elements are able to adapt to new roles when needed.
- *Data* and *information*: essential entities used by advanced systems and organizations to enable goal-driven functionalities through *communication*. Data are raw material for information, extracted from signals by observation – or by technical means in modern society. Information is a generic term for meaningful data that an intelligent system can utilize. *Knowledge* is conceptually organized information, and *language* is the primary means of communicating it – between humans. Intelligent machines are starting to achieve lower levels of knowledge and (natural) language.

Intelligent systems – animals, humans, and advanced machines – are able to create an information-based model, an internal representation of their environment, in order to successfully carry out their tasks. Due to the complexity of reality, such an internal representation can never be complete in full detail. *Overcomplexity* of target objects and environments leads to the need for optimal representations and information processing strategies. We encounter overcomplexity in many forms in the theories and models of this book. Sometimes using a simple model is sufficient although we know the structure and function of the object in detail. Often, at the frontiers of science, the best we know and can do is to use an undercomplex model that is far from perfect but still useful. Overcomplexity may be seen, for example, as statistical behaviour of a system, leading to a probabilistic model. A more powerful model may be able to represent the system deterministically. Many systems are inherently complex, e.g., chaotic, in a way that strongly limits the possibility of modelling and behaviour prediction.

For creatures of biological origin, advanced information processing is the result of continuous evolution through mutation and selection. The use of language and conceptual thinking has helped humans to survive in nature. An integral part of this evolution is information technology and engineering, including books and newspapers, telephone, radio and TV, computers, and mobile communications. Computers and human-engineered technical systems are already superior to humans in strictly logic-based computation tasks. *Pattern recognition* tasks are where humans still strongly outperform machines. It remains to be seen how technology will evolve, and what forms of human and machine information systems will exist in the future.

1.4 About the Rest of this Book

Communication by sound and voice is a many-sided and multidisciplinary topic. Thus, it is not possible to write a single book that covers all related aspects in depth. As discussed above, many subfields of science and technology must be combined to compose an up-to-date view of the whole field. This textbook emerged from a need to cover most of this field rather than concentrating on a single subtopic. More specialized publications can be found when more detailed information is needed.

1.5 Focus of the Book

The title of this book, *Communication Acoustics: An Introduction to Speech, Audio, and Psychoacoustics*, reflects the broad and interdisciplinary approach taken by the authors. It is, in some aspects, quite close to traditional presentations that are entitled something like

Communication Acoustics, which actually is the name of the Finnish book by Matti Karjalainen from which this textbook emerged.

The most central topic of this book is related to auditory perception. Every communication channel consists of a source, a channel, and a receiver. In our case the ultimate receiver is the human auditory system. This is found to be the most challenging link by far in the communication chain. Transmission channels and many sound sources can be formulated by physical and technical models more easily. Human speech or music as sound sources are also quite involved, since a full understanding of these phenomena requires including cognitive aspects too. The receiver, especially the human auditory system, is complex and intricate, even starting from the peripheral parts. Due to this complexity, the auditory aspects in communication by sound and voice have often been omitted or greatly simplified, especially in engineering-oriented textbooks. This simplification or omission is also due to the fact that formal and instrumental knowledge of auditory functions is only emerging.

Another motivation for placing special emphasis on the auditory aspects is that perception of sound is almost always the ultimate reason to study sound, since physically sound carries only little energy and thus has a negligible effect on nature or technical systems. The effects of sound and voice on man, through the hearing organ, are why we pay so much attention to the phenomenon called sound. The importance and emphasis of auditory perception does not mean, however, that other topics and disciplines are less essential in the communication chain. It just means that perception requires special effort to meet its complexity and importance.

The motivation for including basic concepts of physics and signal processing in this book is that, although there exist numerous specialized books on these topics, such knowledge is a prerequisite for understanding the communication chain from the source to a human receiver. The reader is presented with some mathematics and semiformal theory on these topics. If the reader finds them difficult to follow, acquiring at least an intuitive view of these concepts is recommended, at least for the first reading.

The hope is that after reading this book the reader can work on development of techniques based on knowledge of the perceptual mechanisms. This goal is approached by introducing the physical, signal-processing, and psychophysical background to the sound technologies, where a human listener is involved. These technologies involve all audio and speech techniques, certain noise-related techniques, hearing aids, and audiology.

The scope of the book is relatively broad, much broader than can be covered fully in a single volume. Consequently, this means that in the depth dimension of knowledge, this book is shallow, really just scratching the surface of existing literature. References in this book are intended to help the reader to find more specialized sources of information and topics that are omitted from this text.

1.6 Intended Audience

Every book is written for some audience, at least in the minds of the authors themselves. In the case of this book, the targeted audience is expected to be relatively broad and possibly quite inhomogeneous. The approach here is probably more from an engineering sciences point of view than any other. Semiformal presentation of facts, theories, and models, as well as knowledge that may be used for practical applications, comes from this starting point. The authors are convinced that such an approach will become more common in other fields and technical disciplines as well. The modelling paradigm discussed above, especially computational modelling, will become increasingly necessary in order to master such phenomena as auditory perception and modelling of complex communication situations. This is believed to be true both in basic research and in the development of engineering applications.

References

Hunt, F.V. (1992) *Origins of Acoustics*. Acoustical Society of America.
von Helmholtz, H. (1954) *On the Sensation of Tone*. Dover Publications.