Acoustic Communication: The Precedence Effect

Jens Blauert
Institute of Communication Acoustics, Ruhr-University of Bochum, Bochum, Germany, jens.blauert@rub.de

Jonas Braasch
CIRMMT, Department of Music Theory, McGill University, Montreal, Canada, jb@music.mcgill.ca

The acoustic modality is of paramount importance for human inter-individual communication. Consequently, the human auditory system is highly differentiated and able to perform sophisticated tasks such as the identification, recognition and segregation of concurrent sound sources in acoustically adverse conditions – e.g., in reverberant or noisy environments. To this end the different stages of the system at peripheral, sub-cortical and cortical level act in a coordinated manner. In this lecture we take the auditory Precedence Effect as an example to discuss the role of the different stages of the auditory system in complex sound-localisation tasks. Further, we consider different strategies of modelling auditory functions.

1 Introduction

In spaces with sound-reflecting walls, the sound which a source emits arrives at a potential receiver via different paths: (i) There is the direct path which, usually, corresponds to the shortest distance between source and receiver. (ii) There are multiple reflective paths. The sounds along these paths change their directions whenever a reflective surface is hit. This means that they arrive at the receiver from different directions and with different delays – as compared to the direct sound. The auditory system is capable of identifying the direct sound in these situations and, within a certain range of conditions, forms a homogeneous auditory event at a direction corresponding to the position of the sound source – disregarding the directions of incidence of the reflective sounds. This complex and highly non-linear phenomenon is called auditory Precedence Effect – or Law of the First Wave Front. It has been object of scientific discussion since a long time [e.g., 14, 17, 24].

2 Phenomenology

To demonstrate the Precedence Effect in the laboratory an often-used set-up consists of two loudspeakers at different directions but equal distance with respect to a listener as shown in Fig. 1 [2]. The signal to one loudspeaker (the lag) can be variably delayed with respect to the other one (the lead). For delays below about 1 ms, both loudspeaker signals contribute to the direction of the auditory event (so-called summing localization, which is exploited in stereo- and surround-sound technology). For higher delays there is still only one auditory event, but its position corresponds to the direction of the lead loudspeaker.

Figure 1: Lateral positions of the auditory events for a lead-lag stimulus pair in dependence of the delay time $\tau_l$ between lead and lag

The directional information in the lag-loudspeaker signal is usually disregarded by the auditory system. However, it is important at this point to stress that this disregard only concerns directional information. That a delayed sound is present can actually be sensed well by the auditory system and is perceived as increased loudness, a coloured timbre and an increased spatial extent of the auditory event – as compared to what is heard when solely the lead-speaker signal is presented. The width of the auditory event is minimal for zero delay and increases monotonically with increasing delay. This holds for broad-band signals like impulses, noise, speech and/or music.

For narrow-band signals the Precedence Effect is less distinct. In this case, the auditory event varies periodically with the inverse of the signal’s center frequency in the Precedence-Effect region. With decreasing bandwidth, this variation can extend to the midline
between lead and lag, and even lead to failure of the Precedence Effect [4].

When the delay of the lag signal is further increased, the auditory event finally breaks up into a primary event and an echo. The “echo threshold” varies from 1 ms for short impulses to up to 80 ms for organ music.

It is has been observed that the echo is perceptually shifted toward the direction of the lead, if the delay between lag and lead is only slightly higher than the echo threshold. With increasing delay time, this shift disappears [11]. For delays slightly below echo threshold the auditory event may also shift slightly to the middle.

What holds for the case of only one simulated reflection – as described above – holds as well for multiple-reflection scenarios. Within a wide range of conditions there is indeed only one homogeneous auditory event, namely, in the direction of the sound source – i.e. the direction of the first incoming wave front determines the direction of the auditory event. Obviously, the potency of our auditory system to deal with these situations in such a sophisticated way, is the basis of our competence in localising sound sources in enclosed spaces and other reflective environments [3, 21, 26].

3 Bottom-Up Modelling

A bottom-up model in the sense as used here is a computational algorithm where the output is exclusively controlled by the input signals to it. Such models, which are also called “signal-driven” ones, can serve to model the basic properties of the Precedence Effect. As they reflect auditory functions which are assumed to be physiologically situated at the peripheral and sub-cortical level, their software architecture often mimics the physiological prototypes by using elements as follows. The incoming acoustical signals, being filtered by the external ears, pass through middle-ear models and are then fed into components which mimic the inner ears. Here, they are spectrally decomposed and split up into ear-adequate – so-called “critical” – frequency bands. In each of these bands the analogue signals are converted into trains of neural spikes – eventually after some adequate compression – just as happens in the human sensory (hair) cells. The spike trains, or signals which describe their temporal density, are then delivered to units which compare the signals from the left and right ears with respect to inter-aural level and arrival-time differences [2, 5, 6, 16, 18, 19, 22, 25]. The level differences can be evaluated with an algorithm based on the function of excitation/inhibition (EI) cells, while the temporal-difference evaluation is usually based on excitation/excitation (EE) cells which, as has been confirmed physiologically, are able to provide an estimate of the inter-aural cross-correlation function.

It turns out that a model with this general architecture is capable of modelling the Precedence Effect for impulsive sounds quite precisely if the model includes stages to simulate the non-linear transduction behaviour of the inner hair cells [16]. Some phenomena of the Precedence Effect could also be simulated by simply focusing on the model analysis on the spectral dominance region around 750 Hz, while ignoring the information within the remaining frequency bands [22]. Yet, problems occur when it comes to modelling the effect for ongoing sound – such as noise, speech or music. A solution to this problem has been provided by combining the EE and EI mechanisms and extending the temporal-evaluation process by a “contrasting” feature [5, 6]. The latter can be realized by a mechanism known as contra-lateral inhibition [18, 19] which takes care of that. After the first wave front has come in, sounds from other directions, i.e. with different inter-aural signal differences, are disregarded for a while as regards their directional information. The inhibition rests for a certain time span before fading away – see Fig. 2.

Figure 2: Contra-lateral inhibition on top of a unit which estimates inter-aural cross-correlation [18, 19]

So far the model structure is strictly bottom-up, i.e. signal-driven. However, there is evidence that the time-span for which contra-lateral inhibition stays effective, after it has been triggered by the first wave front, must be made dynamically adjustable. In fact, it may be determined by other factors than just the acoustic input signals to the auditory system, whereby it is not clear yet at which level of the auditory system the temporal behaviour of contra-lateral inhibition is actually controlled, even though cells showing phenomena related to the Precedence Effect have been previously found [7] in the central nucleus of the inferior colliculus (IC). We shall come back to this issue in Section 6.

4 Auditory Scenes

An important feature of the auditory system is its ability to analyse auditory scenes and identify and segregate different auditory components in them [e.g., 3].
The question arises of whether such components, once they have been identified and segregated, may show divergent behaviour with respect to the Precedence Effect. Actually, there is some evidence that this is indeed the case.

For example, it has been shown in laboratory experiments that with two trumpets playing simultaneously but with different times (a slow and a fast melody line), different echo thresholds can be measured for the two lines [23]. Yet, there is interference between the two lines when in mixture since, being played in isolation they show shorter echo threshold each. As regards the interaction of the two auditory streams, listening test with spectrally-filter noise signals have revealed that the interaction is dependent on frequency. The strongest effect shows up for low-frequencies where the individual echo threshold of a component may increase by up to 40% when presented in mixture [23].

5 Build-Up & Break-Down

The Precedence Effect shows a build-up behaviour as follows. If triggered by a stimulus pair consisting of only one impulsive sound and one reflection, the echo-threshold is shorter than when the stimulus pair is repeated a number of times. This build-up phenomenon has been subject of a number of studies recently [e.g., 1, 8–13, 25]. The main observations are as follows.

The build-up requires a number of trigger stimuli. It is not so much the length of the time interval during which the stimuli are presented but rather the number of trigger stimuli in a row. Ongoing sounds like noise, music or speech do also provide this series of stimuli and, thus, give rise to Precedence-Effect build-up. In terms of the model depicted in Fig. 2, one may thus assume that the time span during which the contra-lateral inhibition is active increases after repetitive stimulus presentation.

It has been experienced that the build-up sequence starts anew whenever the spatial situation varies, e.g., when direct sound and/or reflection(s) change their directions of incidence. For example, when in a two-loudspeaker arrangement as described in Section 2 the lead and lag signals are suddenly interchanged, an echo which had become inaudible after build-up may be audible again for a while, until a new build-up sequence has been completed – an illustrative observation known as the Clifton effect [8] in the field.

At first instance the Clifton effect had been interpreted as evidence for a break-down of the Precedence Effect. Particularly, the understanding was that the Precedence Effect breaks down whenever an auditory scene changes, especially when the change is implausible to the listener. This seems to be reasonable since a such a case it may be of advantage to the listener to analyze the auditory scene again from scratch [8–10, 15, 25].

More recent investigation have, however, revealed a new view of this situation which puts doubt on the break-down hypothesis [11–13]. Clear evidence for the new view can, e.g., be gained by the following action in the Clifton-effect experimental set-up. Shortly after the inversion of the signals, namely, when the built-up for the new situation is in progress or has just been finished, the signals are inverted again, i.e. the initial situation is restored. One will find that the “old” Precedence Effect is still active with a prolonged (built-up) echo threshold. This means that the Precedence Effect for the old situation did actually not break-down but, nevertheless, a build-up for the new situation took place at the instant that this new situation was installed. Further listening experiments were performed in an auditory virtual-environment set-up, as it allows for instant changes of auditory scenes. Among other things, it was found in this way that the built-up effect as observed with a two-loudspeaker set-up also takes place in more complex scenarios, e.g., when switching between differently shaped rooms with multiple reflections. As a result, the Precedence-Effect build-up is currently understood as follows.

When exposed to an auditory scene, the auditory system discriminates between the direct sound and its single or multiple reflections. Consequently the directional information of the reflection is suppressed, i.e. not used when forming the dominant position, i.e. the “centre of gravity” of the auditory event. When the sound source emits repetitive stimuli, such as a train of impulses or ongoing sounds like speech or music, the amount of suppression increases until a saturation. This can, for example, be measured as an increase of the echo-threshold (build-up of precedence). The build-up of precedence happens specifically for the particular spatial situation, i.e. particular directions of incidence of direct sound and reflections. If these directions change considerably, the prolongation of the echo threshold does not hold any more. Instead, a new build-up takes place for the new situation.

Figure 3A depicts how the echoes become audible again when the position of the echoes are changed – or the positions of lead and lag are exchanged as in Clifton’s original experimental set-up. When the position of the lag is switched back to the original position after having it presented from another position once, the echo threshold for the old lead-lag configuration is still sufficiently build up to keep the echo inaudible – see Fig. 3B.

It is a noteworthy feature that a built-up which has happened for a specific situation stays on up to about 9 seconds. This means that if one switches back to this situation within the time span mentioned, the build-up is still effective. It has been confirmed experimentally.
that more than one of these built-up scenarios can exist in the auditory “mind” concurrently – certainly two, but may be even more. Switching between them means that the Precedence Effect is already built-up and, hence, the prolonged echo threshold applies instantly.

Figure 3: Build-up of the Precedence Effect. The echo threshold for a directional channel increases with the number of lead/lag pairs presented in fixed positions. The echo becomes inaudible, if the echo threshold – symbolized by a broken line – exceeds the level of the lag – which is represented by the height of the bars. For a new set of positions, i.e. a further directional channel, a new built-up process is initiated.

6 Top-Down Modelling

As has already been mentioned in Section 3, strictly bottom-up models are not be sufficient to explain all features of the Precedence Effect, e.g., its situation-specific build-up and the capability of having more than one built-up precedence scenario in our auditory “mind” at a time. Consequently, more recent modelling approaches try to include top-down (hypothesis-driven) components. In this way it becomes possible to implement explicit knowledge into the system and, also, to consider multi-modal cues [3]. Such an amendment to our models is of general importance for our efforts in modelling acoustic communication at large.

The idea is that the central nervous system sets up hypotheses based on its situational knowledge and, then, puts these hypotheses up for acceptance or rejection. The output of the model is, consequently, no longer exclusively controlled by its input.

7 Conclusion

The Precedence Effect turns out to be an interesting effect not only due to its immense practical relevance as a fundament of human sound localisation, but also as a subject of scientific research in the context of acoustic communication. Recent results on the Prece-

dence Effect touch upon the nucleus of a prominent general issue of current research in audition and the technology related to it, namely, the question of what are the relative roles of the more peripherals stages of the auditory system (peripheral and sub-cortical stages) as compared to the central stages (cortex). There is a general feeling that the role of cognition in acoustic communication needs further exploration. The Prece-
dence Effect is a good example to support this view.

Figure 4: Architecture of a model of the auditory system which includes top-down (hypothesis-driven) components

References


