Audio Engineering Society
Convention Paper 6819
Presented at the 120th Convention
2006 May 20–23 Paris, France

Designing a Spatial Audio Attribute Listener Training System for Optimal Transfer

Rafael Kassier[1], Tim Brookes[2], and Francis Rumsey[3]
Institute of Sound Recording, University of Surrey, Guildford, GU2 7XH, UK
[1] raf@rafaelkassier.com, [2] t.brookes@surrey.ac.uk, [3] f.rumsey@surrey.ac.uk

ABSTRACT

Interest in spatial audio has increased due to the availability of multichannel reproduction systems for the home and car. Various timbral ear training systems have been presented, but relatively little work has been carried out into training in spatial attributes of reproduced sound. To demonstrate that such a training system is truly useful, it is necessary to show that learned skills are transferable to different settings. Issues relating to the transfer of training are examined; a recent study conducted by the authors is discussed in relation to the level of transfer shown by participants, and a new study is proposed that is aimed to optimise the transfer of training to different environments.

1. INTRODUCTION & BACKGROUND

Interest in spatial audio has increased due to the availability of multichannel reproduction systems for the home and car. Despite various timbral ear training systems having been presented [1-6], relatively little work has been carried out into training in spatial attributes of reproduced sound.

Perhaps the greatest strides in this direction have been taken by Neher [7]. Neher performed a pilot experiment into listener training for spatial audio attributes, and produced a series of unidimensionally varying spatial audio attribute stimulus sets. He argued that in order to train listeners in the perception of spatial audio attributes, one must be able to exemplify changes in specific attributes in the auditory modality. Neher's pilot experiment involved just 5 listeners, but the results indicated that training in detection of differences between, and ranking of, spatial audio attributes can benefit listener performance using the *same* set of stimuli.

It was hypothesised by the current authors that participation in a listener training programme concerned with the spatial aspects of sound reproduction would also help to create listeners that are more consistent and sensitive when evaluating spatial changes in audio reproduction using a *different* set of stimuli to those used in training (so called *transfer of training*). In order to demonstrate its usefulness outside the context of the stimuli used in training, any training scheme would need to show that learned skills were *transferable*.

In documents such as [8], the terms *training* and *familiarisation* (where the procedures involved in listening tests are explained to, and practised by, the test subjects) are used interchangeably, and in [9] *training* could be better described as *practicing* the task. For this research, *training* refers to a separate process where skills are taught and practised in a context not necessarily identical to the test conditions.

According to Shaw and Gaines [10] ambiguities can result when different words are used to describe the same phenomenon, the same words are used to describe different phenomena and different words are used to describe different phenomena. This can result in confusion as to what is meant by one person and understood by another. Care must therefore be taken when selecting appropriate terms to use in the description of spatial audio phenomena.

The first concern addressed by this research was the need for a spatial audio description language that could be used as the framework within which to base the training system. This description language needed to conform to various criteria, such as the need for unambiguous terms that did not overlap conceptually with one another. The resulting *Simplified Scene-Based Paradigm* was published in [11]. Once this framework had been established, a pilot experiment was conducted to examine whether or not participation in a training programme (based upon Neher's experimentation and research) would enhance the consistency, sensitivity and fluency with which subjects performed in a spatial audio evaluation task [12].

The pilot experiment is briefly discussed here, along with a discussion of *transfer of training* and planned future work.

2. PILOT STUDY

Further information about the pilot study is published in [12]. A summary follows.

2.1. Background to the Pilot Study

The pilot experiment was designed to examine the effect of participation in a training programme on listeners' performance in subjective evaluation tasks using stimuli that were different to those in the training programme.

The overall methodology employed for the experiment is summarised below:

- The selection of an appropriate perceptual task to use as a verification of the training system.
- Recruitment, assessment and streaming of a group of untrained (naïve) listeners, so that two sub-groups have similar performance.
- The training of one sub-group of the listeners using a spatial audio attribute training system
- Verification of a training effect by testing both sub-groups of listeners using the perceptual task.

The perceptual task chosen for the pre- and post-training tests was the rating (on a 0-100 point scale) of the perceived width of various solo instruments or singers recorded using 16 different multichannel microphone arrays. The procedure used to record these stimuli is described in [13].

Sixteen naïve subjects were assessed and separated into two equal-skill groups. One of these groups was trained and then both groups were assessed again. Any changes in the performance of each group between the initial and final assessment was therefore likely to be due to their participation or lack of participation in the training programme.

The training of listeners was via a modified [12] implementation of Neher's pilot spatial audio attribute training system [7]. Performance could be measured by examining the correctness of rank-ordering of the contrived stimuli within the training task, and consistency and sensitivity measures of the grading data. The non-trained subjects provided a control group to show the effect of repeating the task at a later date without the intervening training.

2.2. Pilot Study Results

The training system used was shown to be effective at improving performance on the training tasks (involving the training stimuli). There was a statistically significant reduction in the subjects' "wrongness of rank ordering" scores ($z=2.524$, $p < 0.05$) after training [12].

However, the trained sub-group did not show a marked performance improvement over the untrained sub-group of listeners during the ecologically valid tasks (those involving the multiple microphone recordings). A noteworthy training effect, however, was that subjects who had participated in the training tended to use more of the 0-100 point scale in the ecologically valid tests than they had before the training. Conversely, non-trained subjects actually used less of the 0-100 point scale during the second set of ecologically valid trials than during the first.

2.3. Pilot Study Summary

A pilot study was undertaken to study the transfer of training undertaken using one set of stimuli to performance in tasks using another. Sixteen listeners were tested and streamed into two groups showing equivalent performance in a spatial audio perception task. One of the groups underwent a formal training programme which trained in the detection of differences in a spatial audio attribute. The other group did not take took part in any additional training. There was an established "correct" order in which to rank the items, so it was therefore possible to measure the correctness of each subject's response. The trained group showed a significant improvement in the way that they ranked the audio stimuli used in the training scheme. Both groups were then retested on the spatial audio attribute rating task. The only training effect observed was in the way the subjects used the 0-100 point scale to rate the items. The trained subjects used significantly more of the scale to express their judgements after training, whereas the non-trained subjects used significantly less of the scale to express their judgements.

The observed lack of transfer of training from the training stimuli and system to the ecologically valid task of rating spatial audio attributes is a central issue in this research. Issues relating to transfer of training and transfer experiment design were further investigated and used to plan a follow-up experiment. These are covered in the next sections.

3. TRANSFER OF TRAINING

If spatial audio listening skills learned through training can be used in situations outside the training context, then a case can be made for their wider applicability.

3.1. Definition of Transfer

Various authors have attempted to describe or define transfer. Ellis [14] and Wittig [15] use very similar, generalised descriptions that could imply transfer between very similar environments. Detterman [16] is clearly concerned with transfer that occurs between two quite different conditions:

"Transfer of training ... describes situations where the learning of one task influences the later acquisition of some other task" [14].

"Transfer of learning means that experience or performance on one task influences performance on some subsequent task" [15].

"If two situations where the same behaviour occurs are obviously different in important ways, interest is in transfer" [16].

Transfer is often subdivided into two categories: *near* transfer and *far* transfer. Clark & Voogel [17] explain that *near* transfer refers to target contexts that are similar to the training setting, whereas *far* transfer is achieved when skills are applied in "very different" [17] contexts to the trained one. Detterman [16] distinguishes *near* transfer as involving identical situations apart from specific differences, from *far* transfer which describes a "continuum of situations progressively more different from the original learning experience" [16]. *Near/far* transfer in the pilot referred to performance in the training and ecologically valid tasks. An example of *near* and *far* transfer can be seen in the pilot experiment. Performance in rank ordering the stimuli used in training can be seen as a task involving *near* transfer for the trained listeners. Grading the ecologically valid stimuli in a consistent and sensitive way involves *far* transfer for the trained listeners.

Whilst Ellis [15] does not make the *near/far* transfer distinction, he draws the distinction between *positive* transfer (which aids the target task), *negative* transfer (which hinders the target task) and zero transfer (which either indicates that no effect has occurred, or that positive transfer has been cancelled by any negative transfer present). It can be said that, for the most part, the training system used in the pilot study resulted in *zero* transfer (because there was no significant performance improvement over the control group).

Transfer can also be considered as *specific* or *general* [16, 18]. *Specific* transfer involves skills from the initial task aiding learning in the target task. *General* transfer is that which occurs not as a result of specific elements in the original task. *General* transfer involves *warm-up* and *learning to learn* [18].

In addition, Detterman [16] also describes the difference between *Surface structure* (similar controls or overall views) and *Deep structure* (similar internal workings) in training studies.

3.2. Transfer of learning experiment designs

Wittig [14] and Hulse *et al* [18], describe a basic transfer experiment design identical to that presented in Ellis [15].

In the simple transfer experiment an experimental group learns a certain task. Thereafter a control group and experimental group perform a target task. The difference in performance between the two groups can therefore be attributed to transfer between the two tasks that the experimental group achieved. It is important for this experimental design that the groups be "equivalent with respect to factors important in learning the tasks" [15] (Hulse *et al* [18] suggest randomly selecting the control and experimental groups).

Ellis [15] goes on to describe four additional transfer test paradigms, all of which are shown in Table 1.

Experiment 1 is not particularly useful for working out exactly what "A" does for "B", as general factors are not controlled. Experiment 2 uses part of the target task to pre-test both groups (allowing for similar ability groups to be assembled) and control for certain specific factors (see "warm-up", below). The experimental group will still undertake more practice than the control group on the whole. Wittig [14] suggests using a filler task that does not have the specific features of the original "A" task as a means of controlling practice. Experiment 3 is supposed to be useful for inter-sensory transfer experiments [15], but this assumes that transfer will be symmetrical from "A" to "B" and from "B" to "A". Experiment 5 uses time intervals between task "A" and "B" in order to investigate temporal issues. Control groups can be created by not providing them with task "A".

Test	Group	Task 1	Task 2	Task 3
1	Experimental		Learn A	Learn B
	Control		(rests)	Learn B
2	Experimental	Pretest B'	Learn A	Learn B
	Control	Pretest B'	(rests)	Learn B
3	Experimental		Learn A	Learn B
	Control		Learn B	Learn A
4	Experimental		Learn A	Learn B
	Control		Learn A	Learn B'
5	Experimental		Learn A	Learn B

Table 1: Showing five particular transfer test types (after Ellis [15]). Task "A" is normally the initial task, task "B" is normally the transfer task.

The pilot study can be classified as an "experiment 2" type, with the exception that the pre-test was actually the entire transfer test taken before any groups were separated and half of them were trained. In order to control for general transfer factors, an optimisation of the previous method could be the introduction of a filler task for some or all of the control subjects.

3.3. Specific factors in transfer

Specific factors are dependent on the nature of the original task and how that affects the transfer task.

3.3.1. Task Similarity

Transfer is aided if the training and transfer tasks are similar. According to Osgoode (1947, cited in [15]), identical conditions maximise transfer. In addition, if responses are the same, positive transfer will result from similar stimuli, negative transfer will result from "antagonistic responses to identical stimuli" [15], but new responses to previous stimuli will not necessarily result in negative transfer. *Near* transfer is therefore easier to achieve than *far* transfer.

3.3.2. Variety of Previous Tasks

A study by Duncan (cited in [15]) showed that a small increase in the variety in original tasks resulted in an increased positive transfer. The increase was largest from 1 to 2 tasks, progressively less from 2-5-10 [15]. One of the problems with the pilot that was reported by the participants was the lack of variety of stimuli.

3.3.3. Time Interval between Tasks

If memorisation is not required, the time-elapsed between training and transfer tasks does not seem to be an issue [15]. As memory is not the focus of this study, it is not expected that time differences between subjects' experimental sessions need to be strictly controlled in the next pilot.

3.3.4. Degree of Original Learning, Intelligence and Motivation

Ellis recommends extensive practice of the original task, as this reduces the chance of negative transfer [15]. He also advises that intelligence and motivation are factors in transfer [15]. Testing for and controlling intelligence and motivation is a challenge with any type of test. By selecting subjects from existing student groupings (such as first year sound recording students at the University of Surrey) should give a certain level of control over such factors.

3.3.5. Stimulus Predifferentiation

"The greater the relevancy of the initial ... responses to the later ... task, the greater the positive transfer expected" [14]. Stimulus predifferentiation takes place when subjects are conditioned to provide a response to a given stimulus. The more relevant the training tasks appear to the transfer tasks, the greater the transfer can be expected to be.

3.3.6. Task Difficulty

Practicing an easier task may sometimes facilitate better performance in a subsequent task than training on the task itself [14]. This may seem counter-intuitive (given that transfer increases with the similarity of the two tasks).

3.4. General factors in transfer

General factors are independent of the nature of the initial task.

3.4.1. Warm Up

Warm-up results from practice and aids learning by allowing the subject to prepare themselves to attend to the stimuli or adjust to the rhythm of the task. It disappears within hours of the trials [18]. If the transfer task is temporally close to the training task, warm-up should be controlled using a "put in time" task (which warms the subject up using non-related stimuli) [14].

A related concept is *fatigue*, which is the "opposite of warm-up" [14]. Too much practice is likely to make subjects unresponsive to learning opportunities. There is therefore a compromise to be made between sufficient practice (section 3.3.4) and *fatigue*.

3.4.2. Learning to Learn

Learning to learn refers to the process where tasks become easier with practice [15]. It is also the process when subjects learn general principles that can be applied to other situations [14]. Learning acquired through practice is more permanent than temporary warm-up exercises [18]. Hulse *et al*. cited a study by Ward (1937) [18] who showed that practice in learning (memorising) different lists of words allowed subjects to more easily learn (memorise) a specific task. From the presented data, subjects seemed to reach an asymptotic level of performance after around five previous repetitions [18]. Bech found the subjects needed between four and eight repetitions to reach an asymptotic performance level during loudspeaker quality evaluations [9].

3.5. Measuring Transfer

In order to measure the amount of transfer that has occurred in an operation, Wittig suggests looking at either absolute or percentage transfer [14]. In absolute transfer a performance measure (for example the number of errors in the task) are directly compared between the groups. In order to standardise these absolute figures, Ellis [15] provides a number of *transfer formulae* that allow the calculation of a percentage of transfer, allowing for positive and negative transfer to be quantified in a standard way (numerators in the equations are switched over if the performance measure is desirably as low as possible – e.g. errors):

Equation 1 [15]: compares absolute performance of experimental (E) and control (C) groups.

$$\frac{E-C}{C} x100 = PercentageTransfer$$

$$\frac{C-E}{C} x100 = PercentageTransfer$$

Equation 2 [15]: compares absolute performance between E, C and the total possible (T). It is potentially useful if you know the total possible grade achievable.

$$\frac{E-C}{T-C} x100 = PercentageTransfer$$

$$\frac{C-E}{T-C} x100 = PercentageTransfer$$

Equation 3 [15]: compares absolute performance of experimental (E) and control (C) groups. This equation always has a range of -100% to +100%.

$$\frac{E-C}{E+C} x100 = PercentageTransfer$$

$$\frac{C-E}{E+C} x100 = PercentageTransfer$$

Wittig [14] also recommends testing over extended periods of time in order to "catch" any transfer effects that did not show up in the initial transfer tests.

3.6. Teaching for transfer

According to Ellis [15], it is possible to teach for transfer by following certain guidelines (adapted from Ellis [15]).

1. Train and test for specific outcomes - devise the training and transfer tasks so the skills are practiced in a realistic environment that is as similar as possible to the original setting.
2. Analyse the important outcomes of the task and teach and test for those.
3. Provide practice in a "real-world" environment, or final task environment.
4. Allow extensive practice of the original task.
5. Provide examples of concepts and non-concepts in order to demonstrate the applicability of the training.
6. Draw attention to the most important features of the task.
7. Explain general principles in order to facilitate for transfer.

3.7. Procedural and Declarative Objectives and Transfer

Clark & Voogel [17] attempted to explain the many transfer *failures* that have occurred throughout the literature in terms of a confusion between *behaviourist* and *cognitive* procedures.

From their perspective, *near* transfer is limited to specific skills that are not generalisable, and *far* transfer involves *decontextualisation* of skills so that they are widely applicable. They hypothesise that *near* transfer seems to be at the expense of *far* transfer and *vice versa*. They do not expect procedurally trained subjects to be able to easily generalise their skills, and they do not expect those that have generalisable skills to be able to easily use these practically.

They argue that by catering for *near* transfer one is potentially reducing the possibility for the subject to generalise their knowledge, and suggest a number of ways to aid this.

They distinguish between "procedural objectives" (which are useful for near transfer and specify objectives and procedures that need to be mastered), and "declarative objectives" which are more suitable for far transfer (objectives are written in a less rigid manner, allowing more room to experiment.).

Clark & Voogel also suggest using a variety of different contexts for the practice sessions, and the use of analogies as this will help decontextualise the specific skills from the specific stimuli.

They believe that, on the whole, *behaviourist* studies (which tend to foster *near* transfer in their opinion):

- direct and monitor progress
- provide feedback and reinforcement
- test after practice

Whereas cognitive model studies (which tend to foster *far* transfer in their opinion):

- encourage decontextualisation
- encourage discovery
- paraphrase
- use advance organisers
- use analogy
- test the generalisability of learning

They go on to argue that the use of *advance organisers* (explanatory tutorials) in more *cognitive*-based studies can increases *far* transfer further than in more *behaviourist* studies.

3.8. High and Low Road Transfer

Salomon & Perkins [19] attempted to explain transfer in terms of two different phenomena, each being capable of producing flexible skills. They called this *low road transfer* and *high road transfer*.

Low road transfer describes the process by which subjects can learn practices in various situations such that their response becomes automatic. The mechanism for achieving low road transfer is to practice until responses become automatic, and to vary the practice so that new situations are encountered and assimilated by the learner.

The main issue with *high road* transfer is that it involves "mindful abstraction" [19], the decontextualising of the task to allow prior knowledge to help to find a solution.

Unlike Clark & Voogel [17], Salomon & Perkins believe that if one reflects upon *and* practices the behaviour, it is possible that both high and low roads of transfer can be utilised.

3.9. Ahisaar's Reverse Hierarchy and Transfer

Ahissar [20] commented on a paper presented by Wright & Fitzgerald [21], and explained that their experiment showed that transfer for certain learned skills did not occur due to the hierarchical level [22] at which the skills were learned.

The experiment involved the use of inter-aural time differences (ITD) and inter-aural level differences (ILD) to localise sounds on headphones.

Ahissar argued that the learning for each condition must have occurred before the concept of auditory localisation was formed. This has important implications, because the current study is based around the use of higher level *perceptual* or cognitive concepts and not looking at low-level physical parameters. Learning taking place at these higher (fused) levels is expected to be more transferable.

Ahissar also agrees that variety of stimuli is useful in learning, explaining that "initial learning" begins to get the "gist" of the task [20] and that this begins at "generalizing high-level sites". Initial learning appears to be very useful in the quest for optimal transfer of concepts.

3.10. Sternberg & Frensch's Four Mechanisms of Transfer

Sternberg & Frensch [23] suggest four mechanisms of transfer: *Encoding specificity*, *Organisation*, *Discrimination* and *Set*.

Encoding specificity refers to how learning needs to be encoded in the brain in such a way that it is possible to use it in other situations. This can be achieved by explicitly showing students how to apply information and require that they find their own applications for their skills.

Organisation refers to the observation that experts organise learning in a deeper structured level than novices. Organisation can be aided by ensuring that information to be learned is connected logically, either by the trainer or by the trainee.

Discrimination means that information is deemed either relevant or non-relevant for particular situations. If relevant areas are selected for the subject, this will allow them to aid students in choosing relevant objects.

Set relates to having the appropriate mind set required for transfer. Testing for application rather than recall will create a mind set ready to understand concepts rather than facts.

3.11. Transfer on Trial

Detterman [16] conducted a literature survey of transfer studies and came to the opinion that transfer rarely happens, and when it does it is normally because the investigators have specifically explained to the subjects what is needed in order to facilitate transfer. It is worth noting though, that Detterman is almost certainly *only* interested in far transfer.

He gives the following advice:

- Use double-blind procedures, especially with investigations into general transfer.
- Provide a "filler" task for the control group.
- If subjects are told that something will be useful it should not be a surprise when they use it during a subsequent test. The peril is that they might use a trained method in an inappropriate manner.

This last point conflicts with Sternberg & Frensch's advice to assist students as much as possible [23].

3.12. Transfer Summary

Whilst the subject of transfer of training is complicated by many individual theories, some within broader theories of learning, there are a number of points that help answer the questions: how best to optimise transfer of training, and how best to test for it. In order for the training system to show applicability, as wide a transfer as possible is sought.

Firstly, the training task needs to be similar enough to the transfer task to encourage transfer of training from one environment to the other. The tasks also need to have a wide enough variety to allow the subjects to decontextualise the stimuli, facilitating further transfer. Difficulty should be set very low to begin with.

Warm-up and practice effects should be controlled for the non-experimental group, in this case possibly by creating a "timbral" ear training package for them to use instead.

An ecologically-valid task needs to be provided for the transfer task, and transfer needs to be assessed by setting a specific and meaningful goal. This needs to be measured and assessed using Equation 3 to gauge transfer. The test needs to be geared for application, not recall.

Analogies should be used and reflection fostered in order to encourage generalisation of the skills.

Aid should be provided to the students as far as possible in order to assist *encoding specificity*.

4. FUTURE PLANS

In order to gauge the effectiveness of a spatial audio attribute training regime, it will be necessary to compare it against two control groups. Comparison with non-trained subjects will allow the overall training effect to be quantified. Comparing spatial audio attribute training with a group that practices the task (as described in [8] and [9]) will allow any advantages over simple repetition of the task to be demonstrated.

A series of transfer tests must be used in order to gauge the transfer achieved by the training system and the repetitive practice. During the pilot test the emphasis within the training regime was slightly different to the ecologically valid task (the training focussed on discrimination and correct ranking of stimuli, the ecologically valid task required subjects to be consistent and sensitive in their grading of spatial audio attributes). This led to a potential problem that the training and task were not aligned so that it was not possible to test for *near* transfer.

The task used as a pre-test in the follow-up study should also form the basis of testing throughout the training system, and be the task that is practiced through repetition. By comparing the difference in performance between the various groups of subjects during the pre-test and post-test, *near* transfer can be measured for training and practice. It is expected that this task will be a rank ordering task, as this proved to be useful in uncovering training effects in the pilot experiment. A varied set of artificially contrived stimuli should be used that consistently exemplify different levels of a spatial audio attribute. There should be more than one programme item type due to feedback received during the pilot that called for additional stimuli to be used in training.

To test for *far* transfer, three different (transfer) scenarios are planned. Firstly the post-test task will be repeated using a different set of stimuli. Examining the performance (between groups rather than pre-post) in these tasks will therefore indicate how effectively training and practice transfer to stimuli other than those practiced on. Secondly, the testing paradigm will be changed from a ranking task to an attribute rating task, which will verify how the training and practice transferred to different situations (test paradigms). Thirdly the stimuli will be reproduced in a different manner and the original (contrived) stimuli to make the stimuli more ecologically valid. This will result in stimuli where many different attributes of the sound reproduction might change. The ability of subjects to discern and describe a particular sensory characteristic in a "sea" or "fog" of other sensory impressions is more important than sensory acuity [24]. If training or practice could be shown to improve performance with such stimuli (whether using the original or transfer test paradigms) then this would be powerful evidence for their wider usefulness.

Therefore for the next experimental phase, the following hypotheses will be tested:

- Both the trained and practice groups will show improved performance in the initial test and transfer tasks over the untrained group, and over their previous performance. (Because practice and near-transfer training will aid with the initial test).

- The practice group will show improved performance over the trained group for the initial test and stimuli, because they practiced on a task and stimuli closer to the initial task.

- The trained group will show improved performance over the practice group for the other transfer tests. Because more decontextualised training and varied examples will lead to greater *far* trainer.

A full report will follow in a future paper.

5. ACKNOWLEDGEMENTS

The research presented in this paper is supported by a studentship grant from the Engineering and Physical Sciences Research Council (EPSRC).

6. REFERENCES

1. Quesnel, R. and W.R. Woszczyk. *A Computer-Aided System for Timbral Ear Training*. Presented at the *AES 96th Convention.* 1994. Amsterdam: Preprint No: 3856.

2. Quesnel, R. *Timbral Ear Trainer: Adaptive, Interactive Training of Listening Skills for Evaluation of Timbre Differences*. Presented at the *AES 100th Convention*. 1996. Copenhagen: Preprint No: 4241.
3. Miskiewicz, A., *Timbre Solfege: A Course in Technical Listening for Sound Engineers.* Journal of the Audio Engineering Society, 1992. **40**(7/8): pp. 621-625.
4. Letowski, T., *Development of Technical Listening Skills: Timbre Solfeggio.* Journal of the Audio Engineering Society, 1985. **33**(4): pp. 240-244.
5. Olive, S.E. *A Method for Training Listeners and Selecting Program Material for Listening Tests*. Presented at the *AES 97th Convention*. 1994. San Francisco, CA: Preprint No: 3893.
6. Brixen, E.B. *Spectral Ear Training*. Presented at the *AES 94th Convention*. 1993. Berlin: Preprint No: 3474.
7. Neher, T., *Towards A Spatial Ear Trainer*, Doctoral Thesis. *Department of Music & Sound Recording*, University of Surrey, Guildford, 2004.
8. ITU-R BS.1116-1, *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*. 1994-1997, International Telecommunications Union: Geneva, Switzerland.
9. Bech, S., *Selection and Training of Subjects for Listening Tests on Sound-Repoducing Equipment.* Journal of the Audio Engineering Society, 1992. **40**(7): pp. 590-610.
10. Shaw, M.L.G. and B.R. Gaines, *Comparing Conceptual Structures: Consensus, Conflict, Correspondence and Contrast.* Knowledge Acquisition, 1989. **1**(4): pp. 341-363.
11. Kassier, R., T. Brookes, and F. Rumsey. *A Simplified Scene-Based Paradigm for Use in Spatial Audio Listener Training Applications*. Presented at the *AES 117th Convention*. 2004. San Francisco, CA.: Preprint No: 6292.
12. Kassier, R., T. Brookes, and F. Rumsey. *A pilot study into listener training for spatial audio evaluation*. Presented at the *Digital Music Research Network (DMRN) 1st Summer Conference*. 2005. Glasgow.
13. Kassier, R., et al. *An Informal Comparison between Surround-Sound Microphone Techniques*. Presented at the *AES 118th Convention*. 2005. Barcelona, Spain: Preprint No: 6429.
14. Wittig, A.F., *Psychology of Learning*. Schaum's Outline Series, ed. P. Farrell. 1981, New York: McGraw-Hill.
15. Ellis, H., *The Transfer of Learning*. The Critical Issues in Psychology, ed. M.H. Marx. 1965, New York: Macmillan.
16. Detterman, D.K., *The Case for the Prosecution: Transfer as an Epiphenomenon*, in *Transfer on Trial: Intelligence, Cognition, and Instruction*, D.K. Detterman and R.J. Sternberg, Editors. 1993, Ablex: Norwood, NJ. pp. 1-24.
17. Clark, R.E. and A. Voogel, *Transfer of Training Principles for Instructional Design.* Educational Communication and Technology Journal, 1985. **33**(2): pp. 113-123.
18. Hulse, S.H., H. Egeth, and J. Deese, *The Psychology of Learning*. 5th ed. McGraw-Hill Series in Psychology, ed. N. Garmezy. 1980, New York: McGraw-Hill.
19. Salomon, G. and D.N. Perkins, *Rocky Roads to Transfer: Rethinking Mechanisms of a Neglected Phenomenon.* Educational Psychologist, 1989. **24**(2): pp. 113-142.
20. Ahissar, M., *Perceptual training: A tool for both modifying the brain and exploring it.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(21): pp. 11842-11843.
21. Wright, B.A. and M.B. Fitzgerald, *Different patterns of human discrimination learning for two interaural cues to sound-source location.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(21): pp. 12307-12312.
22. Ahissar, M. and S. Hochstein, *The Role of Attention in Learning Simple Visual Tasks*, in *Perceptual Learning*, M. Fahle and T. Poggio, Editors. 2002, The MIT Press: London. pp. 253-272.
23. Sternberg, R.J. and P.A. Frensch, *Mechanisms of Transfer*, in *Transfer on Trial: Intelligence, Cognition, and Instruction*, D.K. Detterman and R.J. Sternberg, Editors. 1993, Ablex: Norwood, NJ. pp. 25-38.
24. Meilgaard, M., G.V. Civille, and B.T. Carr, *Sensory Evaluation Techniques*. Second ed. 1991, Boca Raton, Fl.: CRC Press.