Surrounded by Ear Candy?

The Use of Surround Sound in Oscar-nominated Movies 2000-2012

Svein Høier

Abstract

This article looks at surround sound in contemporary cinema, with the aim of discussing practices of sound design and, more particularly, pinpointing a 'best practice' of surround sound today – focusing here on the practices in the US. The empirical starting point for the analysis is a study of ten Oscar-nominated movies, analysing their soundtracks and especially comparing their stereo and surround versions. The method can be described as a 'directional' listening mode, analysing how the different channels and speakers are used when presenting sonic elements like voices, music, atmospheres and sound effects.

Keywords: film sound, surround sound, immersion, 5.1

Introduction

Surround sound has now moved beyond both the promotional and the experimental stages. Six, seven or eight channels of surround sound will be experienced in most cinemas and in many homes today. Of course, it will always be possible to develop new practices, aesthetics and ways of using surround sound, but this is now definitively an area where one can expect to find a set of conventions, some 'tools of the trade' and a 'best practice'. And today one can also find recommendations from practitioners that reflect this, for instance *Recommendations for Surround Sound Production*, published by The Recording Academy's Producers & Engineers Wing (Ainlay et al. 2004). These recommendations describe many aspects of what practitioners regard as 'best practice' of surround sound in the US today, mostly describing technical aspects but also discussing some of the important aesthetic aspects when designing surround sound for both music and films.

So what kinds of strategies are prominent when sound designers shape voices, music, atmospheric sounds and sound effects in today's surround systems? How do sound designers take advantage of the possibilities such systems present, and how do they overcome the limitations? The discussion of these kinds of questions here is informed by a study of surround sound in American movies, choosing from the soundtracks of movies nominated for an Oscar in one or two of the categories "sound editing" and "sound mixing" for the years 2000-2012, and strategically selecting ten out of a list of 76 films over these 13 years.[1] Diversity in genre and in production year (in the years since the Millennium) have been the two most important criteria for the selection, and

the chosen material consists of action movies like *The Bourne Ultimatum* (Greengrass 2007) and *Skyfall* (Mendes 2012), historical and biographical dramas like *Moneyball* (Miller 2011) and *The Social Network* (Fincher 2010), science fiction films like *Minority Report* (Spielberg 2001) and *Avatar* (Camron 2009), and the adventure *Pirates of the Caribbean: The Curse of the Black Pearl* (Verbinski 2003), along with three films that are perhaps more difficult to categorize: *The Curious Case of Benjamin Button* (Fincher 2008), the war movie *U-571* (Mostow 2000) and *Inception* (Nolan 2010). A detailed analysis of these ten films can give a clear indication of 'best practice' in the American context since the Millennium, and an interest in researching such a 'best practice' motivates the empirical selection.

I have used three different ways of listening in logging and analysing these ten movies. The first involved listening to the surround sound with all the channels activated, to simulate the normal cinema experience (but in a 5.1 system with six speakers and thereby not using the array of speakers used in cinemas). The second mode consisted of listening to the productions in a down-mixed stereo version. This is relevant, because as many will know, there are many situations in which movies are designed for surround sound playback but are experienced in stereo. This will typically be the case when people watch movies broadcast on TV, use computers or watch DVDs or Blu-Rays without a surround system activated. This second way of listening thereby simulates these ways of experiencing the soundtrack in a down-mixed stereo version.[2] The third type involves listening closely to the separate channels in surround, sometimes listening to one channel, sometimes a pair of channels (especially only the two channels used in the back) and sometimes three, four, five or six channels. Overall, this can be described as a 'directional listening mode' that can help to clarify how sounds are spread out through the different channels and speakers when a film is experienced. The direct comparison between listening to a surround version and a down-mixed stereo version is especially important in the following, answering the short – but in some ways also multifaceted and complex – research question concerning the Oscar-nominated films:

> *What do the surround versions have to offer that the stereo versions do not?*

The structure of the following discussion is as follows: the first section will present some possibilities, limitations and problems connected to surround sound on a general level. The second will address the use of voices in the material, after which follows a discussion of the use of sound effects, atmospheric sounds (distinguishing between exterior open space and interior walled space) and music (distinguishing between music with a source within the fictional world (diegetic) and music located outside the fictional world that is represented (non-diegetic music)). The soundtrack is split into these categories for the occasion, to pinpoint how surround is used in different ways in these particular sound categories. The categorization also mimics the main categories that many sound designers use in their technical setup, distinguishing between the mixing groups (or sub-mixes) of dialogue, 'atmos', sound effects and music.

Possibilities, Limitations, Problems – and Pragmatism

When 5.1 surround sound was introduced in the late seventies, audiences were simultaneously introduced to new – and sometimes overwhelming – possibilities for auditory

experiences with sounds coming from all directions, first in the cinema and later, in the nineties, in the home. It is relevant to describe the 360-degree sound experience as the auditory equivalent of spectacle in some ways – as described by Vivian Sobchack when she writes about the promotional trailers for the Dolby Digital format shown in the years around the Millennium (2005: 12). She describes how the "sonic velocity" in some movies can function as the equivalent of – and also equivalent in importance to – the visible 'spectacle'. This 'sonic velocity' seems increasingly relevant today, in both trailers and full length movies of all genres – but is of course most prominent in scenes of high intensity and in the action genre.

Surround sound clearly increased the 'sonic velocity' of soundtracks, and also added what can be described as 'ear candy' – desirable sounds that could heighten an audience's emotional involvement by being intensive, surprising, exaggerated, pleasant, shocking and more. But the introduction of surround sound also involved limitations, as well as a number of possible unwanted effects for an audience. One important limitation is that only a minority of a cinema audience will occupy seats located in (or near) the acoustic 'sweet spot' in the cinema; that is, the best location for listening to the sum of the loudspeakers at once (having the same distance to the various speakers). Even if sound systems in theatres are tweaked to perform at their best, the situation is not optimal, as Kerins describes: "Still, it is simply impossible to make a theatre that will sound equally good (and have the same front to surround balance) heard from a seat on the left edge side of the front row, from the middle of the auditorium, or from the back right corner" (2011: 48). The challenges connected to variation in seating, and possible strategies – like *not* mixing for a sweet spot – are discussed in the recommendations by *The Recording Academy's Producers & Engineers Wing* (Ainlay et al. 2004).

Unwanted effects in surround sound have similarly been discussed by different authors, for instance Chion in the nineties (1994: 84) and more recently Elvemo (2013: 33). So even though the practices around surround sound have been developed and refined over the years, unwanted effects still continue to be troublesome and debated. When problems are described, the 'exit door effect' (sometimes called the 'exit sign effect') is often mentioned as well as, to some degree, the 'in-the-wings-effect'; that is, situations in which the audience's attention can potentially be led away from the narrative and the diegetic space because of the way sound elements are presented at the sides and back of the cinema. In this regard, Tomlinson Holman describes how sound designers need to be careful when using the surround channels: "Called the exit sign effect, drawing attention to the surrounds breaks the suspension of disbelief and brings the listener 'down to earth'" (2008: 116). Other writers have similarly focused on how such effects can be avoided through the thoughtful use of surround (Kerins 2011: 72-74).

Overall, one can say that the combination of possibilities, limitations and possible dysfunctional results of surround sound has resulted in a practice that involves a healthy pragmatism: trying to take advantage of possibilities, coping with limitations and, just as important, *avoiding* dysfunctional results – and presenting a 360-degree sound experience only when it is fruitful to do so. On this basis, many movies do not include all the possibilities presented by surround sound; one example of this is how sound designers use frontal and centred localization of dialogue in almost any production, as discussed in the next section.

Frontal Voices that Demand our Attention

Almost all films present important and dominant voices as frontal and centred – thereby depending highly on the centre channel in a 5.1 surround system. This is also prominent in the material studied in this case. There is some supplementary use of the left and right channels, but these two channels are usually used only to add 'a light touch' of reverberation to centred voices. The laboratory scenes in *Avatar* can exemplify how voices are given such spatial definition by using the left and right channels to add reverb. The dialogue scene in an auditorium in *Inception* is a similar example, and the same effects are also put to use in *Pirates of the Caribbean: The Curse of the Black Pearl*. The material also presents some examples of reflected onscreen dialogue in the back channels, but this strategy is not prominent. One example is *Skyfall,* where one can find some scenes in which this kind of reflected sound is used very lightly to indicate spacious rooms, but it should be noted that most of the dialogue in this particular film is in the centre channel *only*. Overall the dialogue is relatively dry in the ten films, and the acoustic surroundings around voices are similarly very often toned down. Intelligibility is prioritized in the design of almost any dialogue in fiction films, and centred voices will reflect this conventional practice of toning down reflected sound and prioritizing the intelligibility of voices.

The dependence on a frontal voice means that even if surround sound involves the possibility of a 360-degree sound experience – and envelopment – voices are treated traditionally in a surround mix. The dependence on the front channels is also connected to the simple fact that dialogue is mostly presented onscreen – and when it is presented offscreen (often showing a listening character), sound designers will often match the offscreen dialogue with the sound quality of the onscreen dialogue. Another good reason is that diegetic (onscreen) space is mainly experienced as having a depth that is localized *behind* the screen, while the sound speakers are largely limited to presenting the illusion of a sonic space placed *in front* of the screen. In the pragmatic use of frontal voices and the presentation of small amounts of reverb, this will result in voices that are localized at 'screen depth', rather than behind or in front of the screen, and this has proven to work well.

One can say that all voices meant to be interpreted semantically by an audience are placed in the foreground and stay frontal – and mostly depend on the centre channel. It should be noted, however, that when it comes to offscreen dialogue one can find some rare situations in which such dialogue is presented *only* through the use of the surround channels, which happens, for instance, a couple of times in the two movies *U-571* and *The Social Network*. But such examples make up less than one per cent of all the audible dialogue heard in the material chosen here, so it is still fair to say that one can hear all voices – and probably understand any movie – by listening to the centre channel *only*.

So what about more special uses of voices in films, like voiceovers or muffled voices from a distant crowd? The material used in this context is very limited when it comes to examples, but it is very plausible that the use of such voices also will depend on the centre channel, like the voiceover in *The Social Network*. When it comes to distant crowds talking – sometimes called 'walla' in the US and 'rhubarb' in the UK – such atmospheric voices are quite often presented through the use of back channels, but these cases will generally belong to the two next categories of atmospheric sounds.

When there was a minimal use of voices in the promotional trailers for Dolby Digital, THX and other technologies connected to surround sound around the Millennium, this

was probably done for a good reason. The experience of sound effects – and in these promos, ear candy – moving all around an audience is far more suitable for 'showing off' surround sound than voices are. The best practice for a voice is rather a static place-ment in the front and the creation of a merging of voice and character by depending highly on the centre channel. The audience accepts this static centred dialogue because of audio-visual "magnetisation", as described by Chion (1994: 69). This term refers to how voices are experienced as spatially merged with characters, even if they differ in visual and auditory position. The left and right channels, and to some degree the two surround channels, can be used to add a touch of reverb and spatial definition to dialogue, but this is done very sparingly in the ten productions discussed here.

What happens when voices are experienced in the stereo version instead of surround? The difference between the two versions will only be marginal when it comes to voices. When six surround channels are combined in a stereo setup, a voice in the centre channel (in the 5.1 setup) will be distributed equally to the left and right stereo channels, creating a 'phantom centre' and giving the experience of a (physical) centred voice. And when the surround channels are used to present reverberated voices, the experience of (diegetic) space will surely be changed in the stereo version, but not in a very notable way. In sum, the difference will be far more recognizable when *other kinds of sounds elements* are mixed down to stereo, like sound elements within the five categories discussed in the next sections, beginning with sound effects.

Sound Effects that Pull Us Into the Action

Like voices, sound effects are prioritized and placed in the foreground in almost any film, but in contrast to voices, sound effects are used in ways that sometimes take full advantage of surround sound capabilities. This is because voices will generally have a static localization, while sound effects can sometimes be positioned very dynamically. The contribution of sound effects to 'sonic velocity' can, for instance, be very notable when sound effects follow the movements of visual sources in 'three dimensions' (and the trajectories of sources when sources are located offscreen). Sound effects will often be dynamic in this way, and can sometimes create an audience experience of 'a ride'. For instance, the audience can experience riding on the back of James Bond's motorcycle in *Skyfall*, jumping onto the 'cars of the future' in *Minority Report*, flying the colourful creature in *Avatar* or off-piste downhill skiing in *Inception*. In such cases, the sound designers can use sounds of various passing objects to try to pull the audience into the action, and these kinds of 'rides' will again often be combined with some sort of a mov-ing and/or handheld camera.

Another typical use of sound effects that has strong links to moving objects is when the visual side stays more static, while bullets, spaceships, choppers or other objects fly by an audience. Discussing how bullets are presented in *Saving Private Ryan* (Spielberg 1998), sound designer Gary Rydstrom writes: "The movement of these sounds give us the impression of being in the action, having the closest possible proximity to the hor-rors, and of being as unsafe as the men onscreen" (2008: 197).

This kind of moving localization by using a dynamic panning of sound effects is an important strategy when designing surround sound today, and contributes greatly to the experience of 'sonic velocity' in films. But sound effects will sometimes also be used

as "offscreen trash", as described by Chion (1994: 84), using for example the sounds of explosions and crashes that have no *direct* visual reference – and that will never be visualized in a direct way. Chion mentions *Die Hard* as one film that uses such 'offscreen trash' in the nineties, and one will find similar use of sound elements in the more recent Oscar-nominated movies studied in this case. In *U-571*, for instance, explosions from depth charges explode in the surround channels only, presenting the audience with the stressful situation of being in a submarine that is under attack from above and at the same time lacking visual reference.

One possible dysfunctional result when using dynamic panning, or the more static sounds in the case of 'offscreen trash', is the risk of creating the aforementioned 'exit door effect'. Simplified, one can say that abrupt sounds without visual cues can make us 'turn our head'. But such a dysfunctional result can also be avoided – or at least reduced – by depending on context and established links between sound elements and the diegetic world that is presented visually – both 'weak' and 'strong' links. Returning to the example from *U-571*, it is no surprise *that* depth charges are suddenly exploding in the surround channels; what is far more surprising is *the specific moment* this happens. This makes it possible to scare the audience without creating immediate scepticism. This is also relevant in, for instance, apocalyptic war scenes and high intensity car chases – these kinds of scenes make it plausible that offscreen crashes and explosions can happen at almost any time, and such sound design might work very well – even when direct visual cues are lacking.

The most dominant use of sound effects, however, involves elements that are visually presented on the screen, and such onscreen sounds mostly depend on stereo capabilities. In such cases the surround channels will mostly be reserved for presenting a low volume reverberation of highly frontal onscreen sound effects – a use that is similar to how reverberated voices are placed in the surround channels. This can be the case, for instance, when onscreen doors are opened, swords are drawn, or cars drive away. But in contrast to voices, panning between the left and right channels (stereo panning) is occasionally used to follow moving characters and objects moving horizontally across the screen, thereby taking advantage of the capabilities of stereo.

The use of the Low Frequency Effect (LFE) channel in 5.1 systems often adds drama, force, weight and materiality to sound effects and music, and a stereo version will lack this dedicated presentation of low frequencies. What else differentiates the two experiences when it comes to sound effects? First and foremost, the stereo version will of course still include the stereo panning, but it will lack the 360-degree dynamic localization of sound effects. The two sides in this dynamic localization will be kept in the stereo version – the left side stays left and the right side stays right – but offscreen sounds are suddenly limited to being placed inside a stereo width that can be less than 90 degrees wide. There is no doubt that sound effects represent one area where surround sound brings us into the action in a way stereo cannot – by enveloping the audience, by giving us the experience of moving objects coming towards us and by taking us on rides. Similar enveloping capabilities of surround systems are likewise very important for the atmospheric sounds discussed in the next two categories, but atmospheric sounds are usually used in a far more subtle way – and the consequence of down-mixing will correspondingly be less.

Atmospheric Sounds that Bring Us Into the Open (exteriors)

Atmospheric sounds will – like sound effects – often be designed as a 360-degree experience through the combined use of multiple speakers in the presentation of an auditory setting. These kinds of sonic atmospheres can either be based on surround recordings, or be the result of a combination of different mono or stereo recordings. In most cases such atmospheric sounds will involve what Chion calls "passive sounds" – that is, sounds that do not trigger attention around their source, but rather represent territorial sounds that present sonic information in more subtle ways (1994: 85). This can, for instance, be the sound of a noisy city in *Inception*, sounds from various distant human activities in *Skyfall* or the sound of nature in all its variations, like the portrait of the countryside in *Minority Report*. Further, passive sounds are those that are often familiar to us, and can likewise be said to have a comforting quality by being mostly everyday sounds that will very seldom result in the 'exit door effect'.

Outdoor atmospheres will mostly consist of 'direct sound' rather than reverberated sound – connected to the lack of close surfaces that can motivate a noticeable presence of reflected sounds – and will similarly often lack the spatial definition that reverberated sound can produce. But these sound elements are often experienced as distant because of two other important factors, described by Maasø and other scholars: low volume and limited frequency range (in most cases: a reduction of high frequencies) (2008: 37).

When listening closely to the back channels *only* (at high volume), one can sometimes get the feeling that atmospheric sounds can have subliminal qualities, that they are not 'listened to' at a normal volume setting – but are still 'felt'. And *only* listening to the two surround channels in movies like *Moneyball* and *The Social Network* is a very relaxing experience compared to the normal modes of listening to these two films – and films in general. One can also be surprised by what kind of atmospheres are put to use in some films; turning up the volume, one notices that such sound elements sometimes seem a bit random. But this is only audible when one listens to the surround channels only while also increasing the volume quite a bit. When using a normal listening mode – and a typical listening volume – this general 'noise' will probably never result in a raised eyebrow, but will rather envelope us in a subtle and functional way. Atmospheres will, like Holman writes, also contribute to continuity, and "vary from providing a sonic space for the scene to exist in to the practical covering up of presence discontinuities, auditory 'perfume'" (2010: 163).

In general, atmospheric sounds have a more associative connection to the visual side compared to voices and sound effects. In cases of onscreen sounds, the distance to the sources often weakens the connection, seeing for instance a city skyline or a church tower in the background, combining such visual information with correspondent sound elements. The link to the visual side is also 'weakened' because atmospheric sounds are mostly static and are seldom combined with directional panning – and are also not presented with cues that guide a more specific localization in diegetic space. The use of atmospheric sounds therefore contributes to creating a sonic background that often consists of passive offscreen sounds but still gives important information, for instance, regarding the presence or absence of important activity in the offscreen space.

There are important differences between the surround and the stereo versions, even if such atmospheric sounds are mostly very low in volume and stay in the background. When atmospheric sounds are presented as a 360-degree experience, this can result in a

masking of the cinema space, 'replacing' its walls with an outdoor environment. When atmospheric sounds are presented in stereo, this depth will be lacking, and atmospheric sounds will mainly contribute on a more basic level – giving geographical information and contributing to continuity in scenes. The change will be similar when atmospheric sounds are used to presents interior qualities, and are mixed down from surround to stereo – as discussed next.

Atmospheric Sounds that Wrap Interiors Around Us

Listening closely to the Oscar-nominated movies, they often portray interior environments with subtle atmospheric sounds in the background. For instance, there are numerous scenes in different meeting rooms in *The Social Network, Moneyball* and *Skyfall*, and in these scenes the audience is surrounded by such atmospheric sounds as a low hum (an auditory setting created by sounds of ventilation, machinery, noise and more). The sound design can be said to contribute to a possible merging of cinema space and meeting rooms by including such envelopment, and atmospheric sounds can often have this intended result of an audience feeling that they are 'living it' rather than 'observing it'.

In *Moneyball* the interior of a baseball stadium is portrayed by indicating background noise like ventilation systems, machinery, human activity and distant talking. In these interior scenes, the atmospheric sound elements stay passive and are presented as muffled and reverberated sounds that are low in volume and have a limited tonal range. And when James Bond walks through the halls with office workers inside MI6, there is a great deal of distant talking in the background. In these two cases the potential challenge of parallel voices is non-existent, because the audience is guided by huge differences in dynamics, distinguishing the important voices from the background ones, using distant voices to create the impression of two relevant working places by enveloping the audience in an atmosphere presenting background voices.

What happens when such atmospheric sounds are experienced in a down-mixed stereo version and the enveloping effect is removed? In my opinion the change is not drastic, because geographical and narrative information will still be heard. But the stereo version makes it more difficult to 'pull' the audience into the diegetic interior space, similar to an outdoor scene. The competition between sounds will also be increased in the stereo version, and background sounds may be masked entirely by elements placed in the foreground, particularly voices and sound effects. This is because when mixing surround sound one can take advantage of the fact that human hearing has better capabilities for hearing parallel sounds when they arrive from different angles. The possible risk of more nuanced sounds being masked by louder sound elements is also relevant for other sound elements that can sometimes be placed in the background, like diegetic music, discussed in the next section.

(Diegetic) Music that Connects Us with Diegetic Space and Characters

Music is traditionally connected to the diegetic universe by setting the volume, using a limited frequency range and controlling the ratio between direct and reflected sound (the amount of reverb). In surround sound, the use of both static and dynamic panning will similarly influence whether the music heard is understood as diegetic or non-diegetic, by

connecting it with a geographically and spatially defined (diegetic) source or not. This is the case, for instance, in a scene in which two of the main characters in the film *The Social Network* step out of a disco. When the visual focus is changed – through editing and camera work – the diegetic music is dynamically panned and thereby connected to the specific location (the disco).

Diegetic music can, however, also be presented without any specific source, functioning rather as an acousmatic sound that also continues to be acousmatic through the entire scene. In such cases, the implicit connection between music and location will often be helped by the probability of a sound source in the relevant setting, for instance in the opening scene of *The Social Network*. This scene includes diegetic music that seems to be performed live offscreen, and is one example of diegetic music being designed as offscreen 'noise' in the background (mostly presented through the left and right channels in this case). The instrumental music in this opening scene has the characteristics of reflected sound – as if localized just around the corner from the table of the dating couple, but the music does not really indicate any specific direction or geographical location, continuing to be acousmatic and omnipresent during the whole scene.

The difference between diegetic music presented in a surround mix and a stereo mix is not radical, but will – when statically panned – be quite similar to the transformation of atmospheric sounds. However, the difference is more radical when *dynamically* panned diegetic music is presented in a stereo mix, and this transformation will have more in common with how the aforementioned dynamically panned *sound effects* are reduced to moving along the left/right axis in the stereo version. This means, for instance, that the design of music that follows a car with its stereo on, and moves horizontally (left/right), will work well in the stereo version. The change will be more radical when, for instance, the sound of a helicopter arriving in the screen from the back (together with music playing on the helicopter's external speakers) must be reduced to a stereo panning and thereby be 'flattened' in the stereo version. The same flattening can also happen to non-diegetic music in some cases, as discussed in the next section.

(Non-diegetic) Music that Immerses Us Emotionally

Tomlinson Holman describes two options – or sound perspectives – that are relevant in the design of non-diegetic music as a surround experience: 1) Sound designers can simulate some sort of seating in a concert hall, adding reverb to the direct sound when using the surround channels. This is called "direct/ambient" by Holman, and mostly involves direct sound from the front and a reverberated version of the same sound elements from the back. Alternatively, one can 2) simulate what he calls a position "inside the band", thereby creating a more immersive experience whereby instruments (and other sources of musical sounds) can be experienced as localized all around the audience (2008: 87). Both these strategies will include some sort of spatial localization of the audience, either in some sort of a 'performance room' (not necessarily a concert hall) or within the space of the musical performance itself ('inside the band'). The first of these two approaches is absolutely the dominant one in the material studied in this case, but both may be used in different scenes in the same film and can also be combined when mixing the soundtrack.

When Adele sings over the opening credits in *Skyfall*, her voice is positioned frontally and centred, as important voices generally are (like discussed above). The other

channels are mostly used to create a reverberation of her voice. This strategy adds new acoustic definition to the cinema space, and creates a feeling of being surrounded – not by Adele's voice but by walls that reflect her voice back to us – a situation similar to sitting in a concert hall and hearing the singing directly from a stage, but combined with the reflected sounds of singing from the walls and the acoustic environment.

The music Adele sings over is mostly panned between the left and right front channels, other than some bass and strings mixed together with her voice in the centre channel. In the back, the surround channels present a reverberated version of the front, and are toned down compared to the front. So even if surround sound presents possibilities for panning instruments within a full circle, Holman's second category ('inside the band') is not used to the fullest in these opening credits. It is also fair to say that such full circle panning is not a very common experience in cinemas; most film music is highly frontally weighted and the surround channels are most often used for adding space – and spice – rather than for spreading 'the band all around'.

It should be noted that the diffusion of surround formats for music distribution has not been a success so far, and that almost all music recordings today are distributed and experienced in stereo. This also means that audiences and music producers still have stereo as their most important reference for experiencing music – outside the cinema. This will also influence the transformation from 5.1 to stereo, because it will still sound 'normal' (i.e., like stereo) in the down-mixed version. When non-diegetic music also lacks the direct connection to sources within the diegetic space, this will in sum make it reasonable to expect that audiences will not feel they are missing out on something big when listening to the stereo version of non-diegetic music. It should also be noted that when using the back catalogue of popular music, sound designers sometimes only have access to a song's stereo version – something that limits the possibilities to mix such recordings in surround.

What do the Surround Versions have to Offer that the Stereo Versions do not?

When summing up the answers to this question, on a general level the discussion has shown how the surround versions can pull the audience into the diegetic world that is presented by creating envelopment and involvement. The cinema space can potentially be masked out by, for instance, atmospheric and passive sounds that similarly can contribute to the willing suspension of disbelief among members of an audience. The same feeling of envelopment can be created by the use of reverberated voices and reverberated sound effects in the surround channels, creating a relevant acoustic space by reflecting frontal sound elements from the sides and back as well. The use of 'offscreen trash' (active sounds) can add dramatic offscreen actions to the experience – when done with caution. The dynamic panning of sounds – especially sound effects – can further heighten the 'sonic velocity' and present desirable ear candy in films, but surround sound also involves a great deal more than presenting these kinds of ear candy effects in cinemas.

Surround sound increases the possibilities to present more sound elements at the same time, by widening the listening experience to a full circle and by strategically using layering. Voices and sound effects will mostly be placed in the foreground and arranged on top of background layers that most often consist of atmospheric sounds, but

sometimes also diegetic music. Non-diegetic music can again function as either a subtle background element or 'middle ground', or it can dominate the soundtrack entirely by being placed in the foreground. The prioritization between sound elements is followed through in the process of down-mixing to stereo, but sending all sounds through the stereo channels increases the competition between sounds.

The use of a directional listening mode makes it clear that the level of difference between the surround and stereo versions varies within the discussed sound categories. When voices are mixed down to stereo, this has smaller implications than is the case with sound effects. The conservative centring of voices in a surround mix will result in a transformation that is quite unproblematic. Horizontally panned stereo effects transform well, while panning along the front/back axis loses much of its impact. Statically panned atmospheric sounds elements will often function well in the down-mixed version and the change is often rather subtle, but at the same time the envelopment will mostly be lost.

The feeling of acoustic depth does not collapse entirely in the stereo version, because the audience will still get depth cues from the stereo panning, volume settings and frequency range as well as the ratio between direct and reflected sound. Spatial definition will, however, be more 'flattened' and the absence of directional depth cues highly recognizable, even if a number of today's surround practices are continuations of stereo sound practices and depend on the more traditional spatial cues.

The analysis of the ten films revealed some notable variations in strategies and techniques among the films, but the material generally has a large degree of consistency. It is further relevant to note that the findings in this study also resonate well with both early writings on surround, like Chion (1994), along with more recent studies, like Kerins (2011). This sort of consistency among films and theoretical works indicates that it is possible to identify and describe a 'best practice' of surround sound today, and the use of a directional listening mode can help produce empirically grounded knowledge about such a practice. However, it should be noted, of course, that there are obvious limitations in a study of only ten Oscar-nominated films. Thousands of films are released each year, and the possibilities for variation are potentially huge – something that should be thoroughly researched and critically discussed. This study, along with the methods described here, can hopefully contribute to such future discussions.

Notes

1. See, for instance, Wikipedia for a full list: for Best Sound http://en.wikipedia.org/wiki/Academy_Award_for_Best_Sound, and for Best Sound Editing http://en.wikipedia.org/wiki/Academy_Award_for_Best_Sound_Editing.
2. It should be noted that the process of down-mixing can be performed in different ways. Some sound designers take pride in creating the best down-mixed stereo version possible, which will be distributed along with the surround version. But other processes will again involve the automated down-mixing of surround versions to stereo versions using software – and a simple summing of channels. The methods used here do not represent all variations in these processes.

Literature

Ainlay, C., Chiccarelli, J., Clearmountain, B., Filipetti, F., Jones, L. A., Kaplan, R., Levison, J., Ludwig, B., Massenburg, G., Massey, H., Neuberger, H., Ramone, P., Scheiner, E., Schilling, E., Schmitt, A., Skillen, J., & Stubblebine, P. (2004) *The Recording Academy's Producers & Engineers Wing: Recommendations For Surround sound Production*. Available online: http://www.grammy.org/files/pages/SurroundRecommendations.pdf (visited 10.10.2013).

Chion, M. (1994) *Audiovision – Sound on Screen.* New York: Columbia University Press.

Elvemo, J. (2013) 'Spatial perception and diegesis in multi-channel surround cinema', in *The New Soundtrack.* Edinburgh: Edinburgh University Press.

Holman, T. (2008) *Surround Sound – Up and Running*, 2nd Ed. Amsterdam: Focal Press.

Holman, T. (2010) *Sound for Film and Television*, 3rd Ed. Oxon: Focal Press.

Kerins, M. (2011) *Beyond Dolby (Stereo) – Cinema in the digital sound age*. Bloomington: Indiana University Press.

Maasø, A. (2008) 'The Proxemics of the Mediated Voice: An Analytical Framework for Understanding Sound Space in Mediated Talk', in Beck, J. and Grajeda, A. (eds.) *Lowering the Boom: Critical Studies in Film Sound*. Urbana and Chicago: University of Illinois Press, 36-50.

Rydstrom, G. (2008) 'The Use of Surrounds in Saving Private Ryan', (Appendix) in Holman, T. *Surround Sound – Up and Running*, 2nd Ed. Amsterdam: Focal Press.

Sobchack, V. (2005) 'When the Ear Dreams – Dolby Digital and the Imagination of Sound', in *Film Quarterly*, 58/4. Berkeley: University of California Press.

Films and Directors

Camron, J. (2009) *Avatar.* USA: 20th Century Fox, Dune Entertainment, Ingenious Film Partners.

Fincher, D. (2008) *The Curious Case of Benjamin Button.* USA: The Kennedy/Marshall Company.

Fincher, D. (2010) *The Social Network.* USA: Columbia Pictures.

Greengrass, P. (2007*) The Bourne Ultimatum*. USA: Universal Pictures, The Kennedy/Marshall Company, Ludlum.

Mendes, S. (2012) *Skyfall*. Great Britain / USA: MGM.

Miller, B. (2011) *Moneyball*. USA: Scott Rudin, Michael De Luca, Rachael Horovitz.

Mostow, J. (2000*) U-571.* USA: Universal Pictures, Canal+ Image.

Nolan, C. (2010) *Inception.* USA: Warner Bros.

Spielberg, S. (2002) *Minority Report*. USA: Twentieth Century Fox, Dreamworks SKG.

Spielberg, S. (1998) *Saving Private Ryan*. USA: Dreamworks SKG.

Verbinski, G. (2003) *Pirates of the Caribbean: The Curse of the Black Pearl.* USA: Walt Disney Pictures, Jerry Bruckheimer Films.