Synchronization of Audio-Visual Elements in Web Applications

Dale Patterson & Lee Evans
Queensland College of Art
Griffith University
Gold Coast, Queensland, Australia

d.patterson@griffith.edu.au, l.evans@griffith.edu.au

Abstract

The web offers rich visual content, and for many web designers it is the visual domain that dominates their practice. Yet sound, music and audio all play a significant role in the human sensory systems, in some cases sound can be more powerful, especially to emotionally influence viewers. This paper looks at the relationship between audio and visual elements in web content. In particular looking at the power of synchronization of audio and visual content. Bringing together the sensory experience from the two sensory systems to potentially enhance the users overall experience. The relationship between audio and visual elements, and more importantly their level of synchronization was explored through an experimental trial where users were exposed to web content with differing levels of synchronization between the audio and visual components. User feedback from these trials showed that synchronization played a key role in the content that users selected as their preferred items. From these results, several new principles for the effective design of web based audio-visual elements were developed including linking the nature and complexity of the visual forms, to the nature and complexity of the audio forms, to provide greater synchronization and enhance overall user experience.

Keywords: Web Design, Audio-Visual, Synchronization

1 Introduction

Audio-visual synchronization exists naturally in the real world. For the simple act of speech, the visual movement of lips and face are synchronized to the audio waveforms that are heard as the voice. When two objects collide the visual impact is synchronized with the audio sound of the impact. In many ways the human sensory system is tuned, by thousands of years of evolutionary experience, to expect visual and audio synchronization (Moore 2012). This tuning to synchronized items is actually quite complex, especially taking into account that the pure sensory information is not perfectly synchronized. Taking the example of a voice, the visual sight of the lips moving transfers to the users eyes at the speed of light (300,000km/s) and is detected by the eyes sensory systems immediately. The audio sound wave of the voice travels to the users ears at the slower rate of the speed of sound (0.343km/s), and is then detected through vibrations of the eardrum. In a highly accurate sense, the sensation of "I see lips move" will arrive at the brain well before the sensation of "I hear a voice". Yet the brain is able to re-order these input sensations and link them together to provide the user with a sense of synchronization. This is a very complex task and involves more than simply re-ordering the senses in time, it also takes into account other sensory factors including spatial senses that contribute to the users perception of the sound coming from a specific visual source, or more specifically a location or point in three dimensional space (Freeman et al. 2013, Moore 2012).

Given that the brains processing of audio-visual synchronization is so well developed, and that this is regularly strengthened through real world experiences that reinforce the synchronized audio-visual nature of real world items, it is easy to understand why, the human sensory systems quickly identify items that do not exhibit synchronized audio-visuals. In simple terms humans are finely tuned to expect synchronized audio-visuals and anything that is not synchronized stands out as unnatural (Herbst 1997). This is most directly evident when synchronization is missing or lost.

The most recognizable, and one of the most common, examples of synchronization issues of this type is found in television advertising that has been audibly converted from another language, but is still using the original visual footage of the dialogue. Although the audible words are correct, and if heard in isolation would be effective, when those audible words are joined with visual lip movements that do not match (the lip movements speaking the words in the wrong language or accent), the differences are immediately evident to the viewer. This lack of synchronization reduces the overall effect of the audio-visual piece, to a far lower level of quality and impact, than either of the sub (audio or visual alone) parts would have been (Chen 1998, Summerfield 1992).

As the example shows, the human sensory system is finely tuned to the synchronization of audio and visual content. When filmmakers are engaged in filming live actors there is significant focus on ensuring that dialogue is correctly synchronized to ensure effective communication of the message. Yet in web design, including the use of typography, image, vector, animation, motion graphic, audio and video components, there is little focus on closely ensuring that the synchronization of all of these audio-visual components works effectively.

2 Beyond Synchronization to Synesthesia

Understanding the physiological and psychological mechanisms through which synchronization is perceived and interpreted by the user is important. One interesting extension to the understanding of human perception of audio-visual items comes through an understanding of the condition known as synesthesia. In simple terms synesthesia is a condition where the person perceives the world in a quite different and remarkable way. For a person with synesthesia they will interact with the world in the same way that anyone else would, the difference is that when they are stimulated by a sensory input, that input will trigger more than one sensation in the person. Examples of synesthetic experience might include the person seeing the letter form "D" and also getting a sensation of an intense red colour, or hearing a specific musical note and also getting a sensation of blue fuzzy lines. In this condition the synchronization that occurs steps beyond the normal relationship between items and the sensations they trigger, to bring new, usually un-associated sensations, together (Ward 2013, Cytowic 2002, Harrison & Baron-Cohen 1996).

Whether physiological or psychological in nature, the condition is complex. But from a web design perspective the concept it touches on (of using an input of sensory elements) to trigger an enhanced response, above and beyond simple recognition, is valuable.

Considering the brain, as an extremely complex neural network of links and connections between neural cells, allows us to understand how this synesthetic condition might occur. In non-synesthetic interactions, stimuli trigger brain areas in normal ways, leading to normal recognition and sensation. For synesthetic experiences, a simple additional neural link or pathway can fire new brain areas and sensations from the same sensory stimuli or input. For the patient who suffers from synaesthesia these experiences are automatic and likely caused by neural stimulations and links (Ward 2013, Chambel et al. 2010, Whitelaw 2008).

Putting aside synesthetic experience and presentation as a complex condition, it highlights the fact that the human mind is a very complex system and that sensory inputs can trigger sensations in differing ways. One of the most interesting considerations is whether tightly linking the synchronization of audio-visual elements can alter how they are perceived, and the sensations that they trigger. Keeping in mind that the audio component and the visual component each trigger their own sensation, and that the synchronized audio-visual is a complete and additional item or sensation (remember the lip sync issue for dubbed TV advertisements and how the audio and video can be effective on their own, but together the joined "audio-visual" outcome does not work).

Michel Chion's prominent study on film sound, Audio-Vision (1990), provides some important ideas- some of which are often used as a basis for further thought within audio-visual studies. The most prominent theory is that of Synchresis, which is explained as: *"the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time"* (Chion 1990)

A common example used to explain the effect is the treatment of punches in film, the experience seeming somewhat empty without the overstated sound effects that are commonly associated. The brain recognizes the synchrony between the aural and visual events, and instinctively makes the connection, even if those events do not match real world experience. Based on this concept it is possible for the sound and visuals to be entirely unrelated, other than their synchrony (Chion 1990). It is in this area, of visuals being linked to sound, where obvious links (like those of film dialogue or effects like punches) are absent, that there is potential to build synchrony that will enhance the overall outcome.

The power of certain stimuli to generate sensations is well known, with people using music to stimulate everything from factual recall (rhymes for facts) through to emotion and memory (music reminding of events and places or times). From the perspective of neural brain activity, it is clear that the "music" stimulus triggers the basic sensation of understanding the sound but that more complex links build to the emotional and factual memory responses (Snyder 2000).

This research sought to understand the use of synchronization in audio-visual web based content, and to measure whether the use of stronger synchronization between audio and visual components can alter the way that content was perceived. In brain or neurology terms, this can be viewed as measuring whether the synchronization itself can enhance the sensations provided, and hence make the web content more effective.

3 Audio-Visual Synchronization Experiments

To measure the impact of synchronization on web based audio-visuals an experimental randomized trial was carried out. The trial involved 44 participants, 20 (45.5%) being male and 24 (54.5%) female. The participant's average age was 22.2 years, the youngest being 18 and oldest 34. To understand the nature of the group being studied they were queried regarding their media viewing habits and answered the question "How much time do you spend viewing media?" with values on a five point Likert scale, ranging from "Almost Never" to "Almost Constantly". The results, as shown in Table 1 indicate that the group was representative of young adults 18-30 who are relatively high consumers of online media.

How much time do you spend viewing Media?

Answer Options	Response Percent
Almost never	0.0%
Not a lot	7.0%
An average amount	27.9%
A lot	**44.2%**
Almost constantly	20.9%

Table 1: Media Consumption

Following completion of the preliminary questions the testing process involved an initial set of short guidelines on the test process, followed by the main series of interfaces and questions themselves. There were 11 web interfaces presented, each webpage featured an embedded interface incorporating an audio-visual video clip (ensuring that audio-visual synchronization was maintained) that the participants were instructed to watch. The clips comprised of 4 labelled sections - A, B, C and D (see Figure 1 for an example).

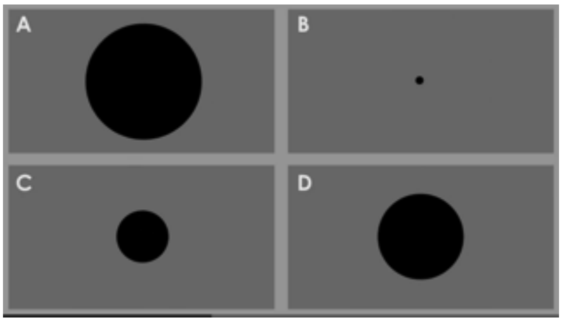


Figure 1: Example of visual display (animated scaling/size)

To avoid any bias caused by recognition of forms, images, motion or sound items of content, the items used as visuals and audio tracks were generic (in the case of Figure 1 a simple circle, an auto generated audio track and auto generated animation – in this case scaling the object over time).

While viewing these items participants were asked 'How complete or correct does each section feel in relation to the audio?' and were instructed to give each section a score on a Likert scale from 1 (worst) to 5 (best). A check-box matrix system was set up to record the users responses. After the main testing process was completed, participants were finally asked if they found the choices difficult, indicating the level of difficulty, before hitting 'done' and completing the test.

Overall the questions provided a mechanism to comparatively measure which visuals (those that were, or were not, synchronized) were most effective. The final questions, regarding difficulty levels, provided an additional qualitative indicator of the users confidence in their decision.

3.1 The Trial Interfaces

To test the importance of audio-visual synchronization, the interfaces that were tested incorporated two main features, firstly an audio track (generated according to the animation/motion type specified) and secondly a set of four visual items (each of the four being identical in form (a simple circle shape, either with or without an overlaid image). These items were then animated according to a particular type of motion or change (see Table 4). The audio track was the same for all 4 visual items (A, B, C & D) but the individual visual items (A, B, C & D) each used a different form of animated change. For one of the items in each visual set the animated change was a perfect match to the change in the audio track (they were synchronized) but for the other three their animated change did not match (they were not synchronised). To ensure the synchronization was maintained (during the web based delivery of the trial) the audio and visuals were combined in the form of web based video/movie file with the audio included, thus avoiding the many asynchronously delivered web scenarios).

Looking at an example, Table 2 shows information about how this example was created, including the audio element, the type of motion being used (in this case a sine wave that is changing the amplitude of the audio track) and the visual element. For the visual component the information describes an image circle with its opacity changing over time. Figure 2 shows the on-screen interface that is presented to the user for this scenario. It shows the four sections A, B, C & D each with a circular image (these are the "Image Circle" described in Table 2) and each image is having its opacity (more or less solid/transparent) changed over time according to the type of animated change (motion envelope) applied (as outlined in Table 3 for this example).

For the example case the user would experience an interface where the audio sound would be a tone whose amplitude (volume) was changing according to a "sine" wave (ie. A steady smooth up and down pattern, see Table 4), they would visually see in section A an image circle with its transparency/opacity changing in a steady up/down cycle that matches the "sine" wave of the audio (this is the synchronized item). Section B would have its transparency/opacity going up and down but with a sharper up/down pattern as defined by the "saw" motion pattern. Section C's transparency/opacity would be changing with a "steady drop and restart" (linear fall) and section Ds opacity would be randomly changing.

Audio envelope:	Sine
Aural form:	Tone
Aural motion:	Amplitude
Visual Form:	Image Circle
Visual motion:	Opacity

Table 2: Descriptor for an example interfaces settings

Visual Section	Motion Envelope Used
A	Sine
B	Saw
C	Fall Linear
D	Random

Table 3: The type of animated change for each visual section

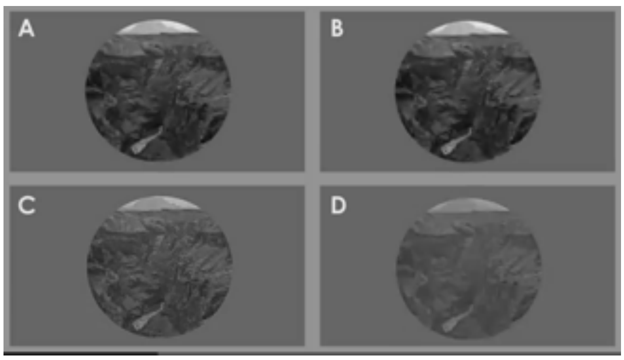


Figure 2: An example implementation (change opacity)

The user would then give a rating for each of the four visuals in terms of how complete or correct that section feels.

For each trial interface different visual features were animated, for example some interfaces use change in opacity/transparency (see Figure 2), others animate movement in X or Y planes (see Figure 3), others animate change in scale or size (see Figure 1), while others animate change in level of detail (see Figure 4). Each of these changes was implemented differently for each of the A,B,C & D sections, and the way they animate was based on one of the set of possible motion types, with only one of those motion types matching the audio motion type used in this audio-visual test (see Table 4).

By selecting just one audio type for each interface and ensuring that one of the visual sections (one out of A, B, C & D in that matching interface) also used the same motion type, there was one synchronized pair in every set while the others were randomly allocated animation techniques from the set in Table 4.

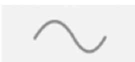
Motion Type	Description	Image
Rise Linear	Steady rising motion, rising linearly from 0-100% in one cycle	
Linear Fall	Steady falling motion, falling linearly from 100-0% in one cycle	
Rise Curve	Rising motion, curved to ease dynamically in and out of motion	
Saw	A linear rising and falling motion. Cycles smoothly	
Sine	Mathematical function creates smooth repetitive cycle	
Random	Erratic, random generated movement	

Table 4: Motion/Animation methods used by Audio or Visual elements

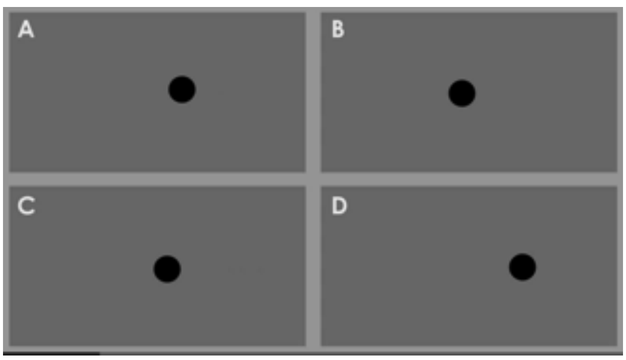


Figure 3: Animation by changing X position over time (moving)

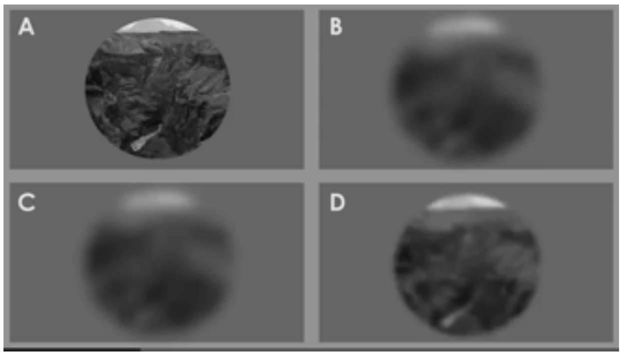


Figure 4: Animation by changing level of detail (blurring) over time

Overall, the trial was designed to provide a mechanism to analyse the users perception of differing levels of synchronization between the audio and visual components of the web page. By delivering these test interfaces as web pages viewed through the normal web browsing systems of the user, this kept the single variable that was being altered to the level of synchronization between the audio and visual elements and allowed for analysis to determine if synchronization was important in how users perceive the audio-visual elements in an online web environment.

4 Results

At the start of the experimental trial each participant was asked “"Do you sometimes experience any visual associations when listening to audio, or audio associations when viewing visuals?", the results of this question are shown in Table 5. This question was asked to identify if there were any participants whose results could bias the larger sample due to existing perceptions.

Do you sometimes experience any visual associations when listening to audio, or audio associations when viewing visuals?	
Answer Options	**Response Percent**
Not at all	4.5%
Not particularly	22.7%
Somewhat	**43.2%**
Strongly	29.5%

Table 5: Existing experience

For each visual component (A,B,C,D) the users were asked to give a rating, in terms of how complete and correct it felt, on a five point Likert scale, ranging from 1 (low) to 5 (high). Results from the interface trials indicated that for 8 of the 11 trial web interfaces, the users selected the synchronized visual component as the most complete and correct.

Across the 11 interfaces with 44 visually animated components (of which 11 were synchronized to the paired audio clips) the overall average rating for the synchronized components was 3.91 (of a possible 5) with the non-synchronized components averaging 2.80.

For the components that were synchronized, and rated the highest (the 8 of 11 cases), there were notable consistencies in their data. Looking at the case of interface 4 (see Figure 2 & Table 6) the audio used a sine wave to over time alter the basic audio tone clips amplitude (volume). The visuals used an image circle with alterations over time made to the opacity of the image based on Saw (A), Sine (B), Fall Linear (C) and Random (D) changes.

Web Item	Motion Envelope Used (changed item)	User Rating					
		1	2	3	4	5	Mean
Audio Tone	**Sine (amplitude)**						
Image Circle A	Saw (opacity)	9%	24%	24%	12%	**30%**	**3.3**
Image Circle B	**Sine (opacity)**	0%	9%	24%	18%	**49%**	**4.1**
Image Circle C	Fall Linear (opacity)	18%	**30%**	24%	18%	9%	**2.7**
Image Circle D	Random (opacity)	**39%**	27%	15%	9%	9%	**2.2**

Table 6: Interface 4 Results

As table 6 shows the synchronized item (B) was rated significantly higher, at 4.1, than all of the other, poorly synchronized, options. Interestingly the second highest rated item was the "Saw" based component (Item A at 3.3). The interesting factor here is the fact that the Saw waveform is the most similar in structure to the 'Sine" waveform. Both produce a rolling up/down pattern, the only difference is the smooth nature of the "Sine" compared to the sharp nature of the "Saw".

This pattern of the synchronized item being most highly rated, followed by the form that was most similar being second most highly rated was observed in all (8 of 11) of the example interfaces where the synchronized item rated most highly. This indicates that being perfectly synchronized is best, however, if perfect synchronization is not possible, the nearer to synchronization the item is, the better the user response will be.

Forms that were particularly poorly synchronized to the audio envelope received notably negative averages in most tests. This was often the case for the "Random" form, its erratic movement not comparing well with many of the other smoother and slower forms (in fact "Random was regularly the lowest rated in tests where it was not the synchronized item). However when the audio envelope was also set to a "Random" form the results showed strong preference towards the match, with the other options receiving consistently low scores in that example. This may indicate that the more erratic and fast movements (found in the "Random" form) tend to adhere easily and well to one another, but are rejected strongly if paired with other slower forms of motion.

Although the majority of the interfaces strongly supported the principle of synchronization leading to better outcomes, the results did reveal some exceptions and factors that appeared to overpower the general concept in some situations. Two of the three interfaces where this occurred involved the use of the "drumbeat" audio clip with a change in the spectrum of the audio clip. In both of these "drumbeat" interfaces the results showed strong preference to the "Random" motion, despite the audio envelopes being a "Saw" and a "Linear Rise", respectively (see example in Table 7).

Web Item	Motion Envelope Used (changed item)	User Rating					
		1	2	3	4	5	Mean
Audio **Drumbeat**	**Saw (spectrum)**						
Black Circle A	Rise Linear (x position)	**55%**	17%	14%	10%	3%	**1.9**
Black Circle B	**Saw (x position)**	7%	14%	**52%**	14%	14%	**3.1**
Black Circle C	Sine (x position)	10%	**35%**	28%	24%	3%	**2.8**
Black Circle D	Random (x position)	**3%**	10%	10%	31%	**45%**	**4.0**

Table 7: Interface 8 Results

These results were a contrast to the usual trend of the matching audio envelopes receiving the highest preference. This seems to show that the motion and rhythmic qualities of the "drumbeat" had overpowered the motion in the spectral change, the rhythm being foregrounded in one's perception and interpreted as closer to the "Random" motion than any of the other options available.

This could relate to the concepts of synchresis in action, the similar erratic motion of the random form being intuitively connected to the fast motion of the drum beat – the prominence of the drumbeat overpowering the less obvious synchronized change in spectrum.

The differing visual forms and "visual changes" used in the study covered a broad spectrum of types including vector and image content involved in movement, resizing, change in opacity and level of detail or blurring (see Figure 5 for examples). These were kept simple to allow the trials focus to be tightly on the synchronization variable, but they were selected to be representative of the common types of visual content and movement found on webpages.

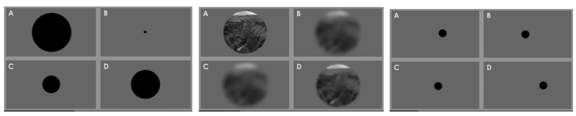


Figure 5: Visual Items and Animation

Across the trial, the users responses were consistent independent of the visual content of the test interface. In terms of user ratings, there was no discernable difference between images and vector content. Both performed with surprising similarity as demonstrated by the example of interfaces 4 & 2. Interface 4 used an image circle with a "Sine" altered audio track (see Table 6 for details) where Interface 2 used a black (vector) circle also with a "Sine" altered audio track (see Table 8 for details). As the results from tables 6 and 8 show, both had the synchronized "Sine" visual highest at 4.1 and 4.0 respectively, followed by the "Saw" visual at 3.3 and 3.4 with the "Random" at 2.1 and 2.3 being the lowest rating. In simple terms the "visual item" had no significant effect on the users perception and rating of that interface.

Web Item	Motion Envelope Used (changed item)	User Rating 1	2	3	4	5	Mean
Audio Tone	**Sine (pitch)**						
Black Circle A	Rise Linear (y position)	21%	6%	24%	24%	27%	**3.0**
Black Circle B	**Sine (y position)**	0%	6%	24%	38%	32%	**4.0**
Black Circle C	Saw (y position)	9%	24%	12%	32%	24%	**3.4**
Black Circle D	Random (y position)	44%	21%	15%	6%	15%	**2.3**

Table 8: Interface 2 Results

This finding indicated that the visual content or technique used, was not as important as the fact that there was a visual change, that was synchronized with the audible change. This ties in with the theory of synchresis that the items do not need to be logically related to benefit from synchronization.

The final element in the testing process involved a final question that queried the users level of comfort and confidence in their choices.

Did you find the choices difficult?	
Answer Options	**Response Percent**
Very much so	0.0%
Somewhat	26.9%
Not Particularly	**53.8%**
Not at all	19.2%

Table 9: User Comfort

User confidence, as shown in Table 9, indicates that the users found the decisions comparatively simple and that they were comfortable with selecting the answers.

Overall the trial results broadly support the theory that the synchronization of audio-visual elements on webpages, whether the individual audio and visual components relate to each other in any other way or not, leads to enhanced user experience.

5 Analysis & Applications

Acknowledgement of the importance of synchronisation in audio-visual content is not new. In fact, in film it has been heavily used since the earliest examples of recorded motion pictures (Fahlenbrach 2005, Vernallis 2004). The concept of synchresis, or using synchronized elements to enhance or create a greater outcome through synchronization of items (whether those items match real experience or not) has also been widely used in film to enhance storytelling and dramatic effect (Chion 1990, Whitelaw 2008). The findings of this study indicate that a similar technique, of intentionally synchronizing audio and visual components (that had no particular link to each other beyond being synchronized) in a web environment, generated interfaces where users rated their interaction with those synchronized systems as significantly better, in terms of how complete and correct they felt. The study identified that this principle applied to a broad range of visual types of content (from simple vector shapes through to images) and showed that the animation of different visual elements (from an items position, scale, opacity to detail) all functioned in the same manner.

In order to isolate the single variable, synchronization, the interfaces were comparatively simple, with 4 main visual components and one audio component. When compared to the rich content of a modern webpage/application (one that may even include 3D and other content types), this limited detail may have an impact on how significant the application of synchronization would be in wider real world application and needs further experimental studies.

The results from the trial showed that perfect synchronization generated improved interfaces, but in addition to this the results showed that components that were not perfectly synchronized, but were close in nature and form to the paired example ("Saw" and "Sine" were good examples of this), performed better than those with lower levels of synchronization.

Interfaces with a lack of synchronization had a notably negative impact on users perception of the interface. This was most evident through the poor performance of the "Random" samples when paired with any form except another "Random". Interestingly when the two "Randoms" were paired, the synchronization between them caused this to perform to a similar level to all other forms of synchronization. Indicating that the effect was not linked to any pattern but instead relates purely to the match of the two items (whether random or structured in nature).

The only key examples where the synchronized items failed to be most highly rated were in the "drumbeat" interfaces. These drumbeat examples highlight the fact that the synchronization needs to be related to a prominent part of the paired audio (or visual) otherwise it may be overpowered by other elements.

For the drumbeat the "Random" pattern was rated most highly, and this can most easily be explained through its faster repetitive pattern, a pattern that is a good match to the rhythm of the drumbeat itself. This would appear to have overpowered the other audio components that were being altered. For the web designer this is an important factor to take into account. Understanding the significant items in a webpage, and their sub-attributes that may dominate other features or items, is important to understanding their role in synchronization.

Those significant elements may not always be obvious. The study by Tsay (2013), demonstrated this principle in a live music performance environment. Showing that in a live music performance, with audio and visual components, the assumption would be that the audio performance would outweigh the visual, yet in this study that was not the case. Much like the rhythm of the drum overpowering the other elements, the visuals in the live music example overpowered the audio (Tsay 2013). This again highlights the fact that the overall audio-visual is more than the sum of the audio and visual components. By adding an additional point of focus to the synchronization between those elements, the web page/application can be enhanced.

In terms of the application of this knowledge in web design, it is clear that there is value in pursuing synchronization. To do this effectively the web designer will need to be able to identify the key content items in their page/application, and within those content items identify the key audio or visual features to target for synchronization. In many ways this is much like the work of the filmmaker in applying the concepts of synchresis. To do this, the filmmaker identifies key points in the story and "enhances" them by applying the synchronized techniques of synchresis to the audio and visual items. The main objective of this process is to link them together in a more powerful manner that will enhance the story.

Returning to the example of the treatment of punches in film, the filmmaker highlights the impact of the punch by synchronizing the visual impact of the punch with an audio effect, note that in most cases this is not the sound of an actual fist impact, but instead is a dramatic audio clip. On its own the audio clip would make no sense, but as part of the "punch audio-visual" its synchronization to the visual is combined by the brain to give an enhanced sense of impact. For the web designer the principle is essentially no different. As the results from this study have indicated, by synchronizing a visual to an audio item the new "audio-visual" is perceived to be enhanced. If the web designer carefully chooses the items to "synchronize" in this way they will enhance their web content.

One of the interesting findings from this study related to the fact that the items being used for this "synchronization" did not need to be related in any way. Much as the audio clip used to enhance the film "punch" is not related to an actual punch, the two web items can also be unrelated, linked only by their synchronization and the knowledge that the brain recognizes the synchrony between the aural and visual events, and instinctively makes the connection, even if those events do not match real world experience.

This allows for a variety of types of web content to be used in this manner, including but not limited to motion graphics, images, video and audio. This research did not address the more complex 3D applications, games or interfaces that utilized more complex devices (Jiang et al. 2011, Patterson 2014, 2007, 2003). Further experimental studies are needer to understand if these principles apply in such complex 3D environments.

Overall the application of these principles creates several new challenges for web designers.

Firstly where audio-visual content is to be used in an online web environment the designer must consider the importance of synchronization (and how to deliver the content in a manner that will retain the desired synchronization). This may mean reorganizing web pages to group content items (eg. and audio-visual file rather than separate audio and visual elements) that maintains the desired synchronization.

Secondly within audio-visual items the synchronization needs to be considered, not just for dialogue and speech but for all types of content from simple motion graphics through to film. As this study indicates, synchronization, even of unrelated audio or visual items brings a greater level of user satisfaction and linked with good design this can enhance webpages.

Thirdly, and finally, the entire webpage that is being interacted with, is essentially one large audio-visual item, and as such the synchronization of the combined set of elements needs to be considered to ensure the most effective interface is delivered. This would mean looking at the visual complexity of the page and matching that to the audio complexity of the sound, while also managing the synchronization of the specific key items to ensure the auditory and visual sensory systems are coming together in the most complete and effective manner.

6 Conclusions

The results from this experimental study have demonstrated that web based audio-visual content, that features synchronization between audio and visual components, receives higher ratings from users. These higher ratings appear to be linked to the human brains capacity to link information from separate stimuli and merge them into an "audio-visual" sensation that enhances the overall experience. This is very similar to the concept of synchresis that is applied in filmmaking.

From a design perspective the ability of this principle to take any pair of audio and visual items (even those that are completely unrelated) and by synchronizing them through motion create an "audio-visual" outcome that provides more to the user than the sum of the two original items.

As many web formats and methods are largely asynchronous in nature, this presents a challenge for web designers seeking to actively synchronize audio, visual and other types of content and will require a design approach that considers synchronization as a key feature.

The findings from this experimental study indicate that for web designers to more effectively present and communicate their information to users, they can benefit from applying synchronization to the content in their online pages/applications. The level of effort that is applied to synchronization in film based media, should be applied just as actively to web content including vector, motion graphics, images, typography and animation.

7 References

Chambel, T., Neves, S., Sousa, C., & Francisco, R. (2010): Synesthetic video: hearing colors, seeing sounds. *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 130-133). ACM.

Chen, T., & Rao, R. R. (1998): Audio-visual integration in multimodal communication. *Proceedings of the IEEE*, *86*(5): 837-852.

Chion, M. (1990): Audio-Vision: Sound on screen, Columbia University Press, New York.

Cytowic, R. (2002): Synesthesia: a union of the senses ,MIT Press,Massachuchetts.

Fahlenbrach, K. (2005): Aesthetics and Audiovisual Metaphors in Media Perception. *CLCWeb* 7(4) 4: 1-9.

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., Leff, A. & Driver, J. (2013): Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex*, *49*(10): 2875-2887.

Graber, D. A. (1996). Say it with pictures. The annals of the American academy of political and social science, 85-96.

Harrison, J. and Baron-Cohen, S. (1996). Synaesthesia: Classic and Contemporary Readings, Blackwell, London

Herbst, T. (1997): Dubbing and the Dubbed Text-Style and Cohesión: textual characteristics of a special form of translation. *Text Typology and Translation, Amsterdam & Philadelphia: John Benjamins*, 291-308.

Jiang, L., Guan, C., Zhang, H., Wang, C., & Jiang, B. (2011): Brain computer interface based 3D game for attention training and rehabilitation. *6th IEEE Conference on Industrial Electronics and Applications (ICIEA),* ***6****:*124-127. IEEE.

Moore, B. C. (Ed.). (2012): An introduction to the psychology of hearing. Brill.

Patterson, D. (2014): Using Interactive 3D Gameplay to Make Complex Medical Knowledge More Accessible. *Procedia Computer Science*, 29:354-363.

Patterson, D. (2007): 3D SPACE: Using Depth and Movement for Selection Tasks. *Proceedings of the twelfth international conference on 3D Web Technology* (12):147-155. ACM.

Patterson, D. (2003): 3D Space: special project in advanced computer environments. Ph.D. thesis. Bond University, Gold Coast.

Snyder, B. (2000): Music and memory: An introduction. MIT press.

Summerfield, Q. (1992): Lipreading and audio-visual speech perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 335(1273):71-78.

Tsay, C. J. (2013): Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences*, *110* (36):14580-14585.

Vernallis, C, (2004): Experiencing music video : aesthetics and cultural context, Columbia University Press, New York.

Ward, J. (2013): Synesthesia. *Annual review of psychology*, *64*:49-75.

Whitelaw, M. (2008); Synesthesia and Cross-Modality in Contemporary Audiovisuals, *Senses & Society*. Volume 3, Issue 3:259-276

Whitney, J. (1981): Digital Harmony: On the complementarity of music and visual art. McGraw-Hill, Inc.