

EBU Recommendation R37-2007



The relative timing of the sound and vision components
of a television signal

Status: EBU Recommendation

Geneva
February 2007

The relative timing of the sound and vision components of a television signal

<i>EBU Committee</i>	<i>First Issued</i>	<i>Revised</i>	<i>Re-issued</i>
<i>PMC</i>	<i>1986</i>	<i>1997, 2002, 2006 (twice)</i>	<i>2007</i>

Keywords: Lipsync, AV Synchronization, Timing, Delay, Clapperboard, VT Clock, Slate

In television, a discrepancy between the instant at which an action is seen to take place and the instant at which the corresponding sound is heard can be subjectively disturbing. To avoid this, it is necessary to ensure that differences in the relative timing of the sound and vision components do not exceed an insignificant value. The value of the relative delay that is just perceptible depends on several factors, notably the programme content and the viewing distance. Experience suggests that particular care needs to be taken in HD production. The following recommended maximum values have been established on the basis of subjective tests of the relative delays at which failure of the synchronism between lip movements and speech becomes perceptible to 50% of observers, under the viewing conditions defined in EBU Recommendation R28 [1].

The EBU recommends,

That, whenever possible, Member organisations should take action to minimise any differences (such as those that may exist within camera channels) in the relative timing of the sound and vision components of a television signal. Preferably, this should be applied by automatic correction techniques, at each point at which significant differences are apparent.

The accuracy of A/V synchronisation at each stage should lie within the range of Audio 5 ms early (sound before picture) to 15 ms late (sound after picture).

This may be measured in the electrical domain from concurrent “blip” and “flash” programme header markers, such as may be present on VT clock programme headers. The flash consists of a peak white field for one full frame duration, and the blip consisting of a 1 kHz tone carried at alignment level for the same period.

These should therefore be retained for this measurement purpose, even within file-based programme distribution. Throughout the programme chain, the audio and video data should be maintained coincident “on the tape” (or in the file or data stream) and not advanced or retarded to account for down-stream processing delays.

At any output intended for emission, the difference in the relative timing of the sound and vision components should lie within the overall (end to end) range given in Table 1 below.

Table 1: Limits of the relative timing of the sound component of a television programme relative to the corresponding picture component

Sound before picture	Sound after picture
≤ 40 ms	≤ 60 ms

Bibliography

- [1] EBU Technical Recommendation R28-1997: Subjective assessment method to be normally used for 625 line television pictures