

al Gruppo tematico per la cinematografia di AES/ATIC/SMPTE

I perché di un vero "Standard"

di Fabrizio Calabrese

Il convegno del 19/01/08 presso l'ANICA si è rivelato prezioso per via del fatto che sono stati portati all'attenzione di tutti noi almeno **tre punti fondamentali** della questione:

1)- Molti dei maggiori esperti (p.es. Savina) appaiono non aver pienamente compreso quale sia il reale ordine di grandezza che separa l'ascolto domestico (TV - Home Theatre - PC) da quello nelle sale digitali correttamente dimensionate... tant'è che insistono nel considerare possibile che uno stesso messaggio possa essere adattato per tutti gli impieghi.

2)- Gli operatori non-audio hanno una percezione assai diversa dalla nostra di quello che è "alto volume" o di quel che è "basso volume", come ha evidenziato la seconda replica al mio intervento.

3)- L'introduzione dei "metadata" sta avvenendo più sulla base della loro proposizione come "prodotto", che in risposta a precise (cioè quantificate) esigenze di lavoro e fruizione.

RISOLVERE QUESTI TRE PUNTI FONDAMENTALI VUOL DIRE IMPRIMERE ALL'AUDIO PROFESSIONALE UN DIVERSO E MIGLIORE SVILUPPO PER IL FUTURO: vale dunque la pena di attivare uno sforzo comune in tal senso.

Punto 1)- In un ambiente domestico la dinamica di qualsiasi tipo di segnale audio (voce, effetti, spot, ecc.) deve essere necessariamente compressa entro limiti molto angusti.

Il semplice rumore del traffico esterno può raggiungere 60-70 dB"A" e, d'estate, cioè a finestre aperte, non è difficile rilevare 50-55 dB"A" Leq. in una normale abitazione di città. E' chiaro che qualsiasi dialogo a basso livello (cioè appunto tra i 50 ed i 60 dB"A" Leq.) rischia di essere compromesso da qualsiasi rumore appena maggiore della media, p.es. da un passaggio di una moto nella via sottostante. Lo stesso accade con archi e flauti, che (in una sala da concerto) emettono livelli di questo stesso ordine di grandezza. Nulla di strano che la pratica totalità degli ascoltatori domestici preferisca dialoghi e musica a livelli non inferiori a 60-65 dB"A" Leq. Una breve prova con un fonometro fugherà ogni dubbio.

All'estremo opposto dobbiamo considerare che i segnali audio più comuni (voce, musica, rumori) sono tutti caratterizzati da un rapporto "naturale" tra il livello medio (RMS) ed il livello di Picco (Lineare) dell'ordine di 20 deciBel. Un minimo d'involontaria compressione (p.es. per il *clipping* dell'ampli della TV) può abbassare questo rapporto a 12-16 deciBel, senza che questo comprometta irrimediabilmente la qualità dell'ascolto. Ma di quali livelli stiamo parlando? Ebbene un normale televisore, o una radio, possono avere ampli da 2-3 Watt, con altoparlanti da 87-90 dB/1W/1m: in pratica questo vuol dire un livello di Picco (Lineare) di circa 90 dB all'ascoltatore (a 2 m.). Togliamo i 12-16 dB di fattore di cresta, e come risultato avremo un livello medio (RMS, o Leq.) di $90 - 14 = 76$ deciBel; se aggiungiamo la ponderazione "A" arriveremo a circa **70 dB"A" Leq.**

Dunque l'intervallo dinamico dell'ascolto domestico è talmente limitato (5-10 dB) da non meritare nemmeno di essere denominato tale. Stiamo parlando di un livello relativamente "costante", centrato sul livello tipico del parlato non confidenziale (65-66 dB"A" Leq.).

Il contesto delle sale digitali è completamente diverso: qui il livello di rumore di fondo può essere tranquillamente inferiore a 30-35 dB"A" Leq., tenendo conto del fatto che il pubblico

tende ad essere (purtroppo) assai scarso di numero, peraltro in sale molto assorbenti, mentre gli apparati di ventilazione e condizionamento immettono soltanto basse frequenze, che non mascherano affatto le voci. Ecco dunque spiegato perché i dialoghi a 50-55 dB" A" sono perfettamente udibili, nei cinema di un certo livello.

Il limite superiore della dinamica nelle sale digitali è facile calcolarlo: le potenze di amplificazione raggiungono spesso i 4-6 KWatt RMS (= il doppio, di picco), su diffusori da 100 dB/1W/1m, con ascoltatori tra 10 e 20 metri: siamo ancora poco al di sotto di 120 dB di Picco Lineare, ma se aggiungiamo la sezione Sub (LFE) ci arriviamo agevolmente. Esistono però sale con potenze installate superiori a 20-24 KWatt, dove è facile superare i 125 dB. Se da questo valore (di Picco Lineare) sottraiamo il normale fattore di cresta di 20-24 dB, ecco spiegati i livelli medi di 100-102 dB" A" Leq. che sono da considerare il limite tra quel che viene definito "emozionante" e quel che inizia a generare fastidio.

L'impossibilità di missare una sola colonna sonora, adatta sia alla TV che alle sale digitali, emerge confermata da quest'ordine di grandezze:

| | Livello sonoro per l'ascolto domestico | Livello sonoro nelle sale digitali |
|------------------|--|------------------------------------|
| Voci sussurrate | 60-65 dB" A" Leq. | 50-55 dB" A" Leq. |
| Voci urlate | 70 dB" A" Leq. | 80 dB" A" Leq. |
| Rumori (massimo) | 70 dB" A" Leq. | 100-102 dB" A" Leq. |

Proviamo ora ad immaginare il tipo di regolazione da adottare nel caso di un ipotetico "compressore-limitatore" che permetta di effettuare automaticamente la conversione di una colonna sonora adatta ad una sala digitale, in modo di poterla ascoltare in casa: l'intervallo al di sotto di 60 dB" A" andrebbe compresso di 10 dB (aumentando il livello), come quello compreso tra 70 ed 80 dB" A" (questa volta, però, diminuendo il livello); infine l'intervallo tra 70 e 100 dB" A" dovrebbe essere compresso e limitato fino a 30 dB (in pratica spianato verso il basso). Se rileggete bene, vi accorgete che le due ultime indicazioni non possono che portare ad una resa qualitativa ignobile...

Quanto appena detto vale per il caso di una regolazione "perfetta", vale a dire per un caso su cento... In pratica basterà collocare una delle due (3?) soglie appena più su o giù, per ritrovarsi voci urlate, rumori troppo bassi ed innaturali, dialoghi inintelligibili e così via.

Anche ammettendo di poter comprimere la dinamica di una colonna sonora digitale entro i limiti indicati, resterebbe insoluto il problema della resa timbrica: i rumori delle scene d'azione (a 100 dB) conterranno per lo più basse frequenze, che a 70 dB saranno attenuate dal normale calo specifico di sensibilità che si ha a queste frequenze ed ai bassi livelli. Nulla di strano se la cannonata di una corazzata finirà per suonare come un colpo di pistola...

Punto 2)- Chi è abituato all'ascolto radio-TV considera "piano" un ascolto a meno di 60 dB" A" Leq., e "forte" un ascolto al massimo livello consentito dall'apparecchio TV, cioè i 70 dB" A" Leq., o poco più, prima dedotti. Chi ascolta da un sistema Home-Theatre o dai diffusori di un PC può avere orizzonti ancora più limitati, in materia...

In una sala digitale i 70 dB" A" sono raggiunti dai normali dialoghi, appena animati o in presenza di rumori ambientali, mentre i livelli dei rumori più forti (esplosioni, motori, ecc.) sono **1000 volte maggiori** per quello che riguarda la gamma delle frequenze medie (da cui il livello in dB" A"), ma ancora superiori per quel che concerne il contenuto di basse frequenze. Non a caso le specifiche *Dolby Digital* prevedono -per i subwoofers- la capacità di emettere livelli 10 volte superiori a quelli raggiungibili dai diffusori restroschermo.

Punto 3)- Di per sé, dei *metadata* che permettano di modificare i livelli di soli 4-6 deciBel

servono a ben poco, nel contesto sopra illustrato: l'unica loro possibile utilità sarebbe dunque quella di poter essere impiegati per allineare i livelli delle voci, nel caso di un ascolto domestico. Occorre assolutamente avere ben presente questo fatto e metterlo bene a conoscenza di tutti gli operatori, in modo di evitare una falsa partenza su un binario morto.

A mio parere, si è creata una certa confusione tra i concetti di "normalizzazione" (da cui il "dialnorm") e quello di allineamento dei livelli (da cui i "metadata"): il problema è che nel primo caso s'interviene in blocco, su tutto il brano, che viene allineato, rispetto al livello d'ascolto o rispetto al livello di Picco Lineare raggiungibile dall'apparato di riproduzione. Nel secondo caso, invece, si vanno ad operare interventi mirati, programmati in funzione di precise direttive della regia o del fonico.

Il problema nasce quando si alternano in sequenza dei segmenti con soli dialoghi ed altri con rumori anche fortissimi... Che si fa? Si ricalibra tutto ogni qualche secondo?

Soprattutto: ma siete sicuri che programmare tutta una sequenza di "dialnorm" e di "metadata" sia un lavoro più semplice e comprensibile, per un fonico, rispetto a quello di assemblare due diversi missaggi (uno per uso domestico, l'altro per le sale digitali)? Io credo proprio di no.

C'è poi il problema degli errori e delle compatibilità: ve la immaginate una colonna sonora delle più "emozionanti" e dinamiche, compressa quanto necessario ad un ascolto domestico notturno, durante la quale la pressione di un tasto del telecomando attiva per errore la dinamica originale?

Chi è esperto di elettroacustica sa che vi è un altro ostacolo, questa volta insormontabile, che è dato dalla diversa risposta in frequenza dell'orecchio umano ai diversi livelli d'ascolto: "dialnorm" e "metadata" possono pur fare miracoli in termini di semplice trasposizione dei livelli, ma non possono assolutamente intervenire sull'equilibrio timbrico di musica e rumori, che verrebbe stravolto proprio dall'alterazione dei livelli.

Infine

E' il caso di ricordarsi del fatto che i trasduttori attualmente (ed in prospettiva) più comuni, cioè gli altoparlanti di PC, Home-Theatre, telefonini ed autoradio, non possono assolutamente raggiungere i livelli di Picco Lineare di un normale impianto Hi-Fi degli anni '70-'90, per i quali è stata tuttavia missata la maggior parte delle incisioni musicali (vedi il successo delle Yamaha NS-10 come monitor). A salvare gli ascolti domestici (FM) ed in auto hanno provveduto generazioni di compressori-limitatori, specie gli Orban, multibanda. Oggi prevale la tendenza a "massimizzare" i *Files* musicali, con appositi software.

Se questa tendenza dovesse cristallizzarsi, tutto il comparto dell'audio professionale cesserebbe semplicemente di avere alcun senso: per questo dobbiamo assolutamente trovare una via d'uscita, quale può bene essere la diffusione di un "protocollo" che sia non solo e non tanto un semplice strumento per migliorare la qualità delle registrazioni e delle colonne sonore, ma anche e soprattutto il mezzo con cui portare a conoscenza del grande pubblico l'esistenza ed i connotati dell'audio professionale di alto livello. E' possibile che questo impulso si estenda, dal mondo del cinema (digitale) a quello della musica, favorendo un ritorno di fiamma delle discoteche e della vera Hi-Fi, con ovvie implicazioni economiche per gli operatori di settore.

Roma 25 gennaio 2008

Fabrizio Calabrese
336-218374