

人声语言在影视后期中的处理

高媛

浙江广播电视集团

[摘要]在影视作品中，声音是重要的组成部分，语言是传递信息最简单的形式。本文，采用举例、分析、实践等方法通过对声音制作中语言的归纳、编辑、处理及修复等功能进行简述。通过介绍有声语言的拾音方式，和录制手段，再到后期中的内容衔接，去除噪音。体现有声语言制作的方式及效果器处理结果的满意度，使人声语言在影视作品中发挥更好的作用。

[关键词]影视对白；声音处理；人声；降噪

[DOI] 10.12252/j.issn.2096-627X.2021.10.508

引言：语言的特性/功能/特点/作用

语言有旋律和节奏的因素。同样语言在内容上包括人物或角色之间的对白、旁白、独白、解说等等。

在影视艺术作品中，语言是人物或角色之间思想交流，情绪表达，以及表现主题的重要手段之一。同时，能够阐述具体含义的语言在影视艺术的各种元素中具有特殊作用。

影视作品的类型种类繁多，因此决定了影视作品中语言的多样性。我们可以把影视作品按照以下方式分类：影视剧、纪录片、新闻、广告、动画、宣传片、电视节目，和晚会等。

一、语言对白制作流程

在拿到工程文件（一般为OMF或AAF）时，对白剪辑师要对所有同期声进行编辑整理。无论是噪声消除，利用同期备用条替换对白中的片段，整理音轨，平滑声音间的过度，还是为ADR剪辑准备声轨，对白剪辑师都要对对白预混创造出无缝连贯的声轨。

在声音制作中，语言的听感要靠前。保证声音无哨音，无爆破音，不失真，气息声适当，避免近讲效应。同时也要不能过尖，过硬，要有温暖感。在影视声音制作中，还需考虑声画关系，例如采访，演员的对白、旁白、独白等，这类声音要考虑到画面的景别，适量调整声音的响度与均衡。

（1）内容剪接整理分轨

根据不同话筒的拾音特性、音色特点，在后期处理时，一般分为几大类：

1. 以森海塞尔MKH416为例的超心型枪式话筒，一般用于电影广播电视的同期拍摄，卓越的指向性和紧凑型设计，有极高的清晰度，极高的灵敏度和极低的自感噪音，为同期录音拾取带来更好的信噪比。但是由于指向性过于集中，同期录制时可能会出现由于话筒轴心偏离声源造成的声音不实、信噪比过低离轴染色等情况。同时，由于拍摄时多工种共同作业的不可控性，可能在拾取人声时同时拾取到现场摄影脚步，现场掉落声等等的串音。针对以上情况，在后期处理时，挑选出可用的声音素材以呈现最真实的影片内容，针对不可用的声音素材进行修补和替换。

2. 以莱克R400A为例的无线腰包发射系统，该系统一般配备隐藏式全指向性话筒。这套录音系统抗干扰性强，内置的RF频谱分析仪可以用于现场勘测，可以准确无误的找到干扰频率。同时有着出色的音质，小巧的体形和稳定的传输频

率。专有的DSP（数字信号处理器）非常低的失真和优越的信噪比保证了同期拍摄的声音质量。一般用于影视同期拍摄时配合枪式话筒成套使用，它的隐藏性能十分出众，拍摄时可隐藏佩戴于演员胸口处。但是也对应会有一些问题：演员服装过于厚重造成的蹭话筒、衣物遮挡造成的声音高频缺失、清晰度高但缺少现场同期空间感。在后期处理时，将此话筒信号与枪式话筒信号叠用，同时要注意相位问题。

3. 以舒尔AXT100为例的无线数字系统，搭配舒尔WBH53T电容式全指向型头戴话筒，对手机及射频设备具有防射频干扰特性，舒适的耳挂，超纤细的可拆卸话筒杆可放置于左右两边，更方便演员在舞台上的表现，配合完成现场扩声及多轨录制的需求。该系统多用于大型演播室录制，由于兼顾扩声使用，若舞台串音过多，会造成直达声过低、串入舞台反射声过多造成的声音“空”；话筒中低频缺失造成导致音色不饱满的问题。以浙江卫视《王牌对王牌》为例，舞台上10人以上同时佩戴，但真正发出声音的演员只有2-3人，剩余人的话筒信号虽有，但大多数都是串音和不可用素材。在后期处理时，要根据画面挑选出可用的话筒信号，保持空间一致，本底噪声一致。

4. 以舒尔AXT200为例的心型指向性手持话筒，拥有极高的信噪比，可以更好的记录语言，频响丰满。一般用于主持人、朗诵等人声拾取，多用于演播厅录制及晚会的录制。但在实际运用中，非音频专业但主持人或演员对话筒使用不规范、话筒距离声源忽近忽远会造成动态起伏过大，话筒距离声源过近又造成近讲效应的问题。在后期处理时，针对一段声音信号，需剪断调整音频块增益，使响度基本一致。若在音频编辑时忽视这一步骤，终混时强行用大压缩量来控制动态，声音会听起来压缩感过强，缺少呼吸感。

在影视同期拍摄时，一般使用枪式挑杆话筒加隐藏式胸麦无线搭配的传统录音方式，既保证了声音的质感又有了人声的清晰度；以《闪光的乐队》为例，在电视棚内综艺录制时，一般采用手持头戴加隐藏式胸麦方式，保证素材完整；以《跑男》为例，在真人秀录制时，一般采用双发双收的无线拾音系统，为避免无线信号丢失或受到干扰，现场还会配备一名话筒员举杆进行备份录制。

根据以上分类，对声音素材进行分类整理，完成音频编辑工作。例如影视剧中，如果挑杆质量好就以挑杆为主，无线辅助，这样声音更加自然；如果挑杆质量不好就以无线为

主，保证声音素材完整。在保证信噪比的前提下，保证电平基本一致。同时对声音素材进行筛减，（比如现场掉落声，摄影脚步，蹭麦等等）。调整相位，避免镶边，反向抵消。接噪，避免环境声时大时小。

在后期制作中可以对不同类别的话筒信号进行分组处理，保证声音的音色，底噪，空间感一致。Boom挑杆及胸麦无线话筒的叠加，既有了人声的“芯”又有了其他表达的内容，使人声既清晰又丰满。

（2）效果器在后期制作中的运用

在影视声音对白制作中，声音后期制作都要处理一系列挑战，例如在人声语言方面，处理优化的对白、配音和采访等费时费力的基础性工作。多年前人们因为技术手段的限制，很难对同期录制的声音进行修复，最终经常会以配音ADR的方式来解决。iZotope RX问世以来，突破了影视对白制作的局限性，提升了同期录音质量增强了同期内容的使用率，降低了后期配音的成本。

随着iZotopeRX的版本不断升级，在音频修复方面的功能更加全面。2019年底问世的DialogueMatch和2021年的最新版本RX9都提升了音频修复问题的能力。

1. Dialogue Isolate（对白提取）

RX9拥有全新的Dialogue Isolate对白分离算法，该算法可以将干净的人声从混乱的背景噪音中提取出来。处理同期声时。可以将对白提取出来，同时也去掉了除对白以外的噪声，环境声及音乐音效。通过控制对白和噪声的增益，调整分离度来控制对白。其中调整模块有Dialogue gain（对白增益），Noise gain（噪声增益），可以预设几种不同的方案进行Compare（对比）进行调节。

2. Ambience Match环境匹配

影视对白中非常重要的功能，Ambience Match是环境声衔接问题的重要解决方式。过去，ADR补录之后要配上相同环境的环境声，就是在一场或者是相同场景找一条同期声素材其中干净的区域，剪切出一些短暂的环境声，再次混合成一条可用的环境声音。这种做法费时费力，且并不能拼凑出长时间且完整的空气噪声素材。

Ambience Match可以解决这类问题，采样一段干净的环境声，点击操作界面的Learn选项，该软件自动学习并渲染出相似的声音。这样就可以把对白语言中环境声进行填补，保证环境声不断的听感。环境匹配中新增的支持多声道及杜比全景声7.1.2的“复杂模式”可以让专业人士将对白和ADR背景联系起来。

3. De-hum（去除电流声）

新的动态模式—De-hum去嗡嗡声算法，在标准版和高级版中可立即消除嗡嗡声/震动或干扰且不降低保真度。

嗡嗡声通常是由于缺乏适当的电气接地引起的，无论是连接到计算机，电源插座还是电源的外围设备。这也扩展到风扇，空调，电加热和电灯开关。

关于可用的参数：

灵敏度Sensitivity：调整将被消除嗡嗡声的量

频宽Bands：控制动态陷波滤波器的数值，根据输入信号更新其增益以避免ringing。为复杂的嗡嗡声调整到更高值，但会牺牲透明度。

滤波器FilterQ：控制陷波滤波器的带宽。

范围选择Range Selectors：通过定义消除嗡嗡声处理上的上届和下届来设置消除嗡嗡声。

4. De-clip（修复失真）

声音在失真的情况下，波形也往往处于削波状态。我们通常选择High模式。Threshold（门限阈值），指的是声音开始削波的位置。将滑块在波峰或者波谷中进行拉动，选择合适的处理的范围，并且还可以进行增益补偿（makeup gain）避免过多的衰减。

5. De-crackle（修复噼啪声）

De-Click经常被用到，多用于较小的咔嚓声，而De-crackle可以修复爆裂声有爆破音质感的较复杂的噼啪声。但是不好声音只能弱化，并不会完全处理掉。可以通过Quality（品质因数）Strength（处理强度）和Amplitude（偏移程度）几个参数进行调节。

6. De-Ess（齿音消除）

人声中难免会有齿音，在影视作品中经常会有类似的问题。所以在修音时要将明显的嘶声进行弱化处理。对于齿音处理有两种算法，分别为Spectral，Classic。其中Spectral是通过检测到齿音的时候，只降低齿音的频段，而Classic检测到齿音就会全频段衰减，其他部分的声音同样也会被衰减。

6. De-Plosive（修复喷麦）

通过Strength（强度）和Frequency limit（频率限制）Sensitivity（敏感度）几个参数进行调试。强度越高去除的喷麦声音越多，敏感度越高，声音调整的越精细；不过也会对原始声音有一定的损伤；频率滑块可以设置到200Hz，低于200Hz频率就会衰减，一般喷话筒都都在低频范围，正常情况下200Hz就够了。

7. De-Reverb（去除混响）

在影视剧对白，人物采访，或是语言类节目中都会有一定声学问题。房间混响过大，环境湿度过湿，甚至建筑环境反射声过多都会影响语言清晰度。自从有了De-Reverb，完美的解决了录音环境声场过大过空的问题。同样，选择一段混响过大的声音进行采样（learn），采样的信号要有人声（干信号）和除人声外空间混响，这样就可以分析出混响的声音信号。可通过滑动Reduction（衰减量）调节多段压缩器参数调整。

8. De-Rustle（去摩擦声）

在影视同期拍摄中，演员的动作过大或者无线话筒没有佩戴好，没有完全贴合都会产生无线话筒的摩擦声。有四种

不同强度的预设，通过Reduction strength（衰减强度）和Ambience preservation（环境声保留）进行手动调节。通过环境声保留的调节寻找最好的听感。

9. Mouth De-Click（去口水声）

口水声是录音中不可比避免的问题，在之前通常是用剪辑的手法把这些声音扣掉。而Mouth De-Click是通过Sensitivity（敏感度）Frequency skew（频率范围）和滑动Click widening（click的宽度）可降低口水音。

10. Voice De-Noise（人声降噪）

人声降噪是最常用的降噪方式，它有两种降噪模式，Adapitive mode（实时降噪）模式下，可以将后面的音频直接进行自动预处理，在点击监听的时候就可以听到处理后的效果。在Optimize for Dialogue/Music（优化对白或音乐）的模式下，软件会对音频进行分析，保留音乐的特性和声音的细节。同样在降噪过程中，点击Learn进行采样，通过Thershold（门限阈值）和Reduction（衰减量）来调整比

例。

11. Spectral De-Noise（综合频谱降噪）

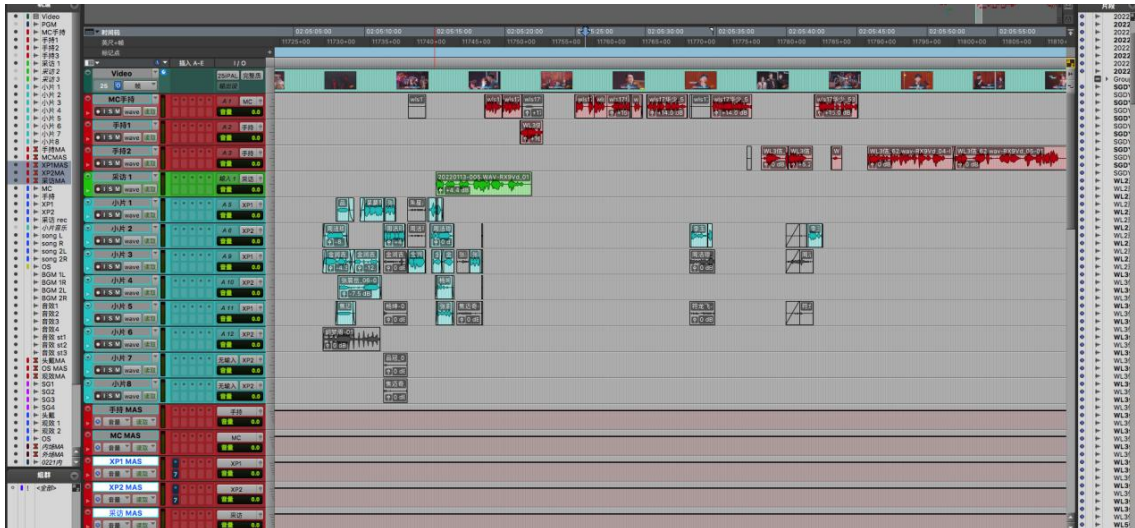
它有两种降噪模式，一种是Learn（学习）采样干净的噪声；另一种是Adaptive mode（实时处理），在播放中将后面的音频自动预处理。在Learn的模式下通常采样同一场景的底噪；而Adaptive mode会随着环境底噪发生变化，比处理不同的底噪，例如场景的切换。在语言节目制作及影视剧中经常使用。

Spectral De-Noise中基本参数也是常用的，比如Threshold（门限阈值），Reduction（衰减值），Quality（品质因数）和Artifact（精确度）等都可以手动调节参数。

二、人声对白制作在实际中的运用

以浙江卫视《闪光的乐队》为例，通过不同类型的对白问题，对应处理方式

以综艺《闪光的乐队》对白工程为例



(1) 按照轨道名称进行不同拾音话筒之间的分配，例如手持话筒是森海塞尔SKM9000系列，拥有卓越的低触碰噪声，更低的口喷灵敏度，以及固定极化电容传感器的声学透性和极高的动态范围。

(2) 我们将不同种类的话筒对应应在不同分轨上，这样在后期处理时更方便整体的量化操作

(3) 调整语言的电平大小，保证其听感一致，不忽大忽小，语言响度控制在-24LKFS

(4) 根据场景不同，进行同场或转场的接噪，除采用同期录制的环境声，也可以在iZotope RX中采用AMB match，平滑过度空气声，保留底噪。

(5) 在遇到喷麦，爆破音，电流声时也可采用不同的降噪手段，使听感达到最佳状态，若同期素材无法使用再进去ADR补录。

(6) 演员角色的频响不同也可在不同分轨上统一做EQ的调整，例如女演员高频过高，为保证舒适的听感可在中高频区域统一对某些频段进行衰减。

(7) 在有些同期拍摄中，没有准备充足的采访中，经常提问的导演声音是申进去的，这种时候，就需要先提高整体电平，再通过降噪的手段进行声音处理以提高信噪比。

规范与标准

国家新闻出版广电总局于2017年12月发布了《中华人民共和国广播电影电视行业标准¹》，其中规定了高清晰度电视成品节目目录制、节目技术质量管控等要求。要求制作时音频采样频率不低于48kHz，采用PCM线性量化，量化比特数不低于16比特。最大真峰值电平不超过-2dB TP，平均响度应为-24LKFS，响度容差为±2LU。同时要求声音效果无异常起伏、明显失真、明显噪声和断点等异常现象。在影视后期对白制作中，要严格按照该标准，并结合自身经验，制作出响度标准、空间统一、清晰度可懂度达标的对白，使人声语言在影视作品中发挥更好的作用。

参考文献

[1] 谭壮满. 文学语言在影视作品中的重要作用[J]. 电视指南, 2017(13): 18-19.