

北京电影学院毕业论文

“类人”角色的特殊语声形象塑造

—— 以《沙丘》中哈克南男爵的对白实验为例

作 者：丁思立

院 系：声音学院

专 业：录音艺术（电影录音）

教学层次：本科

年 级：2019 级

学 号：191061003

指导教师：石宝峰

成 绩：

日 期：2023 年 5 月

独创性声明

本人郑重声明：所呈交的论文是我在指导教师的指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写的研究成果，也不包含为获得北京电影学院或其他教育机构的学位或证书所使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 丁思立 签名日期： 2023年5月14日

关于论文使用授权的说明

本人完全了解北京电影学院有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件及电子文本，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印、或其他复制手段保存论文。

(保密的论文在解密后遵守此规定)

学位论文作者签名： 丁思立 指导教师签名： 王京成
签名日期： 2023年5月14日 签名日期： 2023年5月14日

内容摘要

人类的语声,本就是一种原始而独特的乐器,拥有着最为丰富的色彩表现力。在非现实电影的语声创作中,如果说融合、变形,使其呈现出独一无二的奇特听感需要高度的想象力和创造力,那么创作一种与人类相似的语声则更多依赖理性,它既要在某些方面听感奇特、异于人类,同时又要一定程度上保留人声的基本特质,巧妙的语声设计总能在熟悉与陌生之间找到平衡。

本文的研究聚焦“类人”角色的语声创作,通过分析人类的发声原理、人类体型与频谱特征对应关系的跨学科成果,并结合角色背景、视觉形象、原声带分析、国内外创作经验、原片创作者的思路分享等多方面依据,选定《沙丘》中的哈克南男爵作为“类人”角色语声形象的典型。随后,本文对哈克南男爵的对白部分展开了完整、详细的复现实验,并在此基础上进一步拓展了针对性复合技术手段,初步归纳了VocalSynth、Throat、Dehumaniser、OVox四种主流复合效果器的优劣势和适用方向。最后,笔者综合前文理论分析与系列实验所获经验,对“类人”角色语声创作的总体思路和规律展开了讨论。

结论表明,以人类生理特征与语声特征的对应关系为理论依据,能够有效地指导“类人”角色语声形象的创作。这一结论不仅启示创作者们在预先构思、语声录制、素材分析、手段选择、参数调整等各个步骤都可以参考人声相关理论,也拓展至汉语、英语的语言差异方面。在语言特征和发音习惯的双重影响下,不同语种的母语者发声时存在着细微的生理差异,这一潜在特征应被视为对汉语语声进行针对性处理的重要依据,为中国电影的实际应用提供参考。

关键词: 对白处理; 语音特效; 声音设计; 效果器

ABSTRACT

Shaping Special Speech Sound Representation of “Humanoid” Characters

—— Taking the Experiment of Baron Harkonnen’s Dialogue in *Dune* as an Example

The human voice, originally a primitive and unique instrument with the richest artist capabilities. In the creation of speech sound for non-realistic film , if fusion, distorting or giving them a bizarre and uncommon auditory experience requires a high degree of imagination and creativity, but the voice that is similar to human beings relies more on rationality. It must be peculiar in some aspects, different from human beings, while retaining basic characteristics of the human voice at the same time. Sophisticated speech sound design always finds a balance between familiarity and unfamiliarity.

This research focuses on the speech sound creation of "humanoid" characters. By analyzing the interdisciplinary achievements of human vocal principles, the correspondence between human physique and frequency spectrum characteristics, combining the character background, visual image, original soundtrack analysis, domestic and foreign creative experience, and the thought sharing by the original film creators, Baron Harkonnen in *Dune* was selected as a typical speech sound representation of the "humanoid" character. Subsequently, a complete and detailed reproduction experiment of Baron Harkonnen's dialogues is conducted, and on this basis, the technical means of targeted compounding are further expanded, and the advantages, disadvantages and application tendencies of the four mainstream compounding effectors, VocalSynth, Throat, Dehumaniser and OVox, are initially summarized. Finally, based on theoretical analysis and the experience of a series of experiments, the ideas and rules of speech sound creation for "humanoid" characters were discussed.

The conclusion shows that taking the correspondence between physiological

characteristics and speech sound characteristics of human as a theoretical basis can effectively guide the creation of the speech sound image of "humanoid" characters. This conclusion not only suggests sound designers to refer to human voice-related theories in every step of production, including pre-construction, dialogue recording, sound analysis, methods selection, and adjustment of parameters, but also extends to differences between linguistic differences between Chinese and English. Under the influence of both linguistic characteristics and pronunciation habits, there are potential physiological differences in the vocalization of native speakers of different languages, and the resulting differences in speech sound characteristics may lead to slightly different effects of the same process. The differences in correspondence between the physiological and vocal characteristics of the two languages should be considered as an important basis for complementary processing when targeted adjustments of Chinese speech sounds are needed.

KEY WORDS: Dialogue Processing; Special Sound Effects; Sound Design; Effectors

目 录

中文摘要	I
ABSTRACT	II
目 录	IV
引 言	1
第一章 何为“类人”角色	3
第一节 非现实角色的分类	3
第二节 “类人”角色的概念	5
第三节 “类人”角色的语声形象	5
一. 语声的表义性	5
二. “类人”角色的语声特征	6
第二章 人类的语声	7
第一节 人体发声机理	7
第二节 人声频谱与体型的对应关系	9
第三章 《沙丘》中哈克南男爵的对白复现实验	11
第一节 复现实验思路依据	11
一. 基于角色背景和视觉形象的语声构思	11
二. 对于原声带的分析	12
三. 参考混录师 Ron Bartlett 所述	14
第二节 拟定实验方案	15
一. 录音阶段	15
二. 后期处理阶段	15
第三节 实验过程	16
第四节 实验结果及分析	19
第四章 拓展使用针对性复合技术手段	20
第一节 两种特征的实现效果对比	20
一. 关于“类人”音色特征	20
二. 关于体型特征	21
第二节 处理优势与适用方向	22
第五章 思考与应用	24
第一节 实践中的应用要点	24

第二节 汉语与英语的应用差异	24
一. 关于语言特征	25
二. 关于发音习惯	26
三. 语言综合影响下的声音特征	27
四. 针对汉语的补充调整	27
第三节 “类人” 语声处理思路应用在其他声音中的可能性	27
结 语	29
附 录（一）	30
参考文献	31
致谢语	34
指导教师评语	35

引言

随着电影制作技术和艺术理念的不断发展，视觉特效（VFX）将越来越多色彩纷呈的非现实角色形象呈现在了观众眼前。为了匹配这类角色各异的外型特征，协同视觉一起构建一个独特、有趣、可信的奇幻世界，声音形象的设计必须被纳入进来，而在这之中，角色“语声”正是最关键的部分。

语声（Speech sound），指以言语（Speech）形式传递语言（Language）时所携带的声音信息。这其中不仅包含着与生理学、人体医学相关的声学特征，还包含着与语言学相关的语种特征。作为电影进行艺术表现和情感表达的核心途径，人物语言能够唤起观众直接的心理感受，是与观众连接最为紧密的桥梁。目前，电影理论对此开展的研究大多集中于对白的语义范畴，但实际“声”的部分同样具有涵义，能够揭示出表述的潜在意蕴，并且更加灵动和富于变化。由此，本文引入语声概念，着重探讨“声”如何在电影中实现表义功能。

从早期 R2D2 经典的机器人语声，到近年《猩球崛起》（*Rise of the Planet of the Apes, 2011*）中的凯撒等一类充满野性的语声，非现实角色的银幕表征支撑起了光怪陆离的语声形象，同时也极大地拓展了语声的创作空间。于艺术理念而言，电影声音是艺术，其创作应完全由创作者的主观意识和审美感受驱动，任何循规蹈矩的条框都会成为自由创作的束缚，但对于方法论而言，电影声音是“科学”，绝对主观的判断和普适技巧的挪用无法在每一次都发挥作用。基于这一理念，本文开篇就提出并阐释了非现实角色中“类人”与“动物与异型”两大基础类别划分，而区分这两种角色类别之所以重要，是因为它们从根本上指向的是两种不同的创作思路——人化与兽化。对于兽化程度较高的“动物与异形”角色，我们有着较为丰富的经验参考和历史追溯，其怪异的外型也为语声创作提供了更多的可能性，而对于近似人类的“类人”角色，如何让其非人的同时依旧保留一定的人化特征这一问题，目前仍有较大的研究空间。

“类人”角色语声形象塑造的难点在于：第一，语言的复杂性。语言表述过程中，为准确对应不同的元音，发声者体内每一时刻的生理形态都发生着极其微妙的变化，因此在对素材进行处理时，必然需要顾及这些细节以最大限度地保证语声的流畅度、自然度，这大大限制了创作的手段和尺度；第二，与人类相似的形象通常暗示了与人类相似的生理构造和发声机理。在此背景下，我们无法再依赖于制作怪兽声时常用的大幅度变速变调、融变人声配音与生物吼叫等激进方式，而是需要参考人体发声理论来对其进行更加细微、准确、理性的处理，从而

实现有机与无机、人化与兽化的巧妙平衡。

本文以《沙丘》 (*Dune*, 2021) 中的哈克南男爵一角作为“类人”角色的典型, 旨在通过一系列的具体实验, 以及在此基础上进一步拓展的针对性复合技术手段, 探讨以人类发声原理来指导“类人”语声创作的可能性, 这或许是此类语声创作的最终要义。

第一章 何为“类人”角色

本章目的在于说明题目中“类人”角色定义的来源和相关特征，明确本文的研究范畴。第一节将提出一种适合于声音创作的角色分类方式，并说明分类的背景、依据和具体划分机制；第二节将进一步阐述“类人”角色的特征和次层级划分；第三节引入“语声”概念，进而对此类角色独特的语声特征进行简要分析，并总述其实现思路。

第一节 非现实角色的分类

非现实角色，一般指在艺术作品中与真实人类、生物角色相对的虚拟角色，它们往往与现实中的人类相似又相离，表现出一种异化的魅力。对于这类角色的讨论，从语义学角度，多强调与人类相较的“他者”意涵^①；从视觉造型角度，多强调对原形态的“常态”进行艺术改造和主观处理。而非现实角色的实现要义就在于超越原“常态形”的限制，创造出一种“非常态”的、具有新奇性、怪异性和不可名状视觉特点的新主观形象^②。

外星人、超人、怪兽、怪物、赛博人、机器人……，观众依靠角色的银幕表征总结出了一系列的角色类型，这些名称通俗、直观、贴切，但在类别层级上存在重叠，也无法实际涵盖非现实角色的所有形态。

对于这类角色，目前暂未形成一种受到广泛认同的权威分类方式，并且由于学科特质、创作模式的差异，各领域对其内部的范围划分、层级深度、分类机制也都不尽相同。例如，在电影美术领域，造型设计和服化设计工作有着高度的具象性，因此他们对非现实角色的分类是依据实现手段来确定的，即需要特效化妆完成的角色、需要动作捕捉技术完成的角色两种；在电影理论研究领域，角色的分类通常会与意识形态隐喻联系起来，可分为追寻身份认同和自我投射的“人型”、和承载人类对未知事物恐惧的“异形”^③；而文学领域对于角色的认识则更为主观，习惯于细致刻画每个角色独有的叙事背景和人物特性，往往不对其进行类别性的划分。

电影声音创作在这一点上与文学创作相似，强调为每个角色设计独属于他的

^① 徐竟涵. 异星情境与创世想象：雷德利·斯科特的异形美学[J]. 当代电影, 2021, No.305(08):pp171-176.

^② 高蕾. 异形美学分析[D]. 西安美术学院, 2007.

^③ 徐泽南. 美国外星人电影的意识形态研究[D]. 湖南师范大学, 2020.

声音形象。这样的创作方式精准、细腻，但也更加依托于经验积累，需要创作者们反复摸索才能领悟不同类型角色的处理规律。因此，为了能有更加高效、明晰的分类来对声音创作提供支持，笔者在广泛查阅各领域资料，并结合角色声音创作思路，以及声音与视觉形象的匹配关系后，尝试提出如下分类方式：

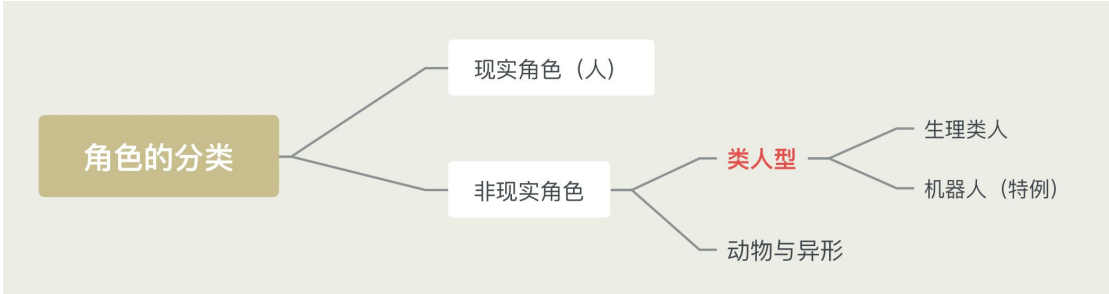


图 1.1.1 角色分类机制

首先，角色的一级类别分为现实角色和非现实角色，其划分条件为是否完全符合人类的“生物”性，符合即为常规的现实角色，不符合即为非现实角色。其次，非现实角色的次级类别分为“类人型”和“动物与异型”，其划分条件有二：第一，是否外观上与人类相似；第二，是否具备完备的语言能力，同时满足两项条件的角色可称为“类人”角色，无法同时满足或均不满足则归为“动物与异型”一类中。例如，外观体貌似人，且具备人类语言能力的普罗米修斯（图 1.1.2）、斯诺克（图 1.1.3）均为“类人”角色，而外观与人差异较大的 ET（图 1.1.4），或并不具备成熟语言能力的格鲁特（图 1.1.5），均为“动物与异形”角色。总体来看，如果说“类人”角色强调的是“人化”，那么“动物与异型”角色则更加注重表现“兽化”特征。



图 1.1.2 普罗米修斯

图 1.1.3 斯诺克

图 1.1.4. ET

图 1.1.5 格鲁特

这种分类机制较为清晰地定义了几种角色的类型及其划分条件，但实际创作中的形象往往更为复杂多变。当界限模糊、难以判定时，应依据角色定位、创作目的、导演诉求等多方面信息综合考虑。

第二节 “类人”角色的概念

由前述分类可知,“类人”角色是指具有与人类相似的外型构造,并具备人类同等级成熟语言能力的角色类型。考虑到本文的研究目标集中于“类人”角色中的一个分支,所以对这一类别内部进行次级划分是必要的。

“类人”角色,主要指“生理类人”型角色,即本文的研究对象。人造人、虚拟人、外星人等等,只要满足两项基本条件,不论存在形式如何,都隶属于这一类别,其声音设计也应符合生理类人的属性,最大限度保留人类特征。但除此之外,“类人”角色的创作还存在另一种倾向——机器人。这类角色处于“机器”与“有机生物体”的临界点,更加强调人工的物理特质。它的身体是机械零部件构成的,存放它意识的大脑也是由程序和算法组成的,因此它所具备的语言能力并非全面,更准确地来说应该是“词法(morphology)”能力。

语言能力是评判生物智能程度的标准,视觉形象是声音设计的重要依据。这样的内部结构和外貌造型共同决定了机器人的声音会更加偏向“无机”,与“生理类人”型角色的创作思路存在一定差异,故此特例不属本文研究范畴。

第三节 “类人”角色的语声形象

一. 语声的表义性

“语声”(Speech Sound)^①是现代信息通信技术中的一种基本概念,指的是人在语言表述时所发出的声音。本文对这一词汇的借用旨在强调除了语言学语境下的“话语”语义外,音色、语气、语调、语速等承载角色特征的语声要素,同样包含着重要的信息^②。

电影声音作为一种表现理念、情感的艺术^③,蕴藏着无数难以言喻的涵义,而语声正是其中重要的组成部分。影视作品中,语声所携带的信息主要分为两类,即人物相关信息和潜在意义信息。其中,表现人物相关信息主要依赖于语声可以传达的三种特征:一为生理特征。包括角色的性别、年龄、体型、健康情况等;二为心理特征。包括角色内心的情感、个性等;三为社会特征。包括时代、地域、文化背景、社会地位等。而对于潜在意义而言,语声表达的能力则主要体现在对

^① How Does the Human Body Produce Voice and Speech?[EB/OL]. USA: National Institutes of Health, 2022.4.12, <https://www.nidcd.nih.gov/news/multimedia/how-does-human-body-produce-voice-and-speech>.

^② 席妍. 论人物语声在影视叙事中的艺术功能[J]. 文艺争鸣, 2018, No.290(09): pp205-208.

^③ 汪月波. 米歇尔·希翁的电影声音理论研究[J]. 当代电影, 2018, No.266(05): pp125-129.

剧作内涵的揭示上。在这一方面，创作者们不仅能够通过不同角色语声特征的对比、呼应，强化叙事的戏剧性，还能够以语声细节的设计来暗示人与人、人与环境、人与自己的关系，甚至隐晦地传递出某种哲学深意。

总的来讲，角色语声呈现出的是一种综合塑造能力，能够协同其他视听元素共同为推动叙事和艺术表达服务。

二、“类人”角色的语声特征

与不具备语言能力，只能发出拟人化、情绪化声调来表现情感的“动物与异型”角色不同，“类人”角色的声音是以人类发声原理为基础，再经一定程度上的异化后形成的，具备双重属性的语声。

在高度异于人类角色的语声创作上，我们有着多样化的发挥空间。可以利用语声中的高噪声含量表现侵略性；利用韵律和间隔表现角色的语态和物种；甚至利用与各异生物声音的融合来表现生命力和兽性。这类角色的语声更倾向于强调“造型意义”，一系列高度夸张、效果性、经艺术化加工的表情性语气词^①都可以作为创作的重要工具。但对于“类人”角色的语声创作而言，语义信息的传递才是核心，需要创作者们更加谨慎地进行处理。心理学理论认为，音色感知的“知觉后效应（perceptual after-effect）”增强了对新的或罕见的听觉对象的敏感性^②，不熟悉的人声音色就是这一理论的首要例证。这告诫我们，在“类人”角色语声创作中，不仅需要让它在某些方面听感奇特、异于人类，同时还要在一定程度上保留人声的基本特质，因为这种似是而非的奇妙感受，只有在熟悉与陌生的矛盾交织下才能诞生。也就是说，若想更好地把握角色“人性”与“兽性”的平衡，人类的语声应被视为“类人”语声创作的重要基础。

^① 刘艳. 科幻类电影非人类角色语音及非现实空间的声音设计特征研究[D]. 南京艺术学院, 2022.

^② Yuyan Wei, Lin Gan, Xiangdong Huang. A Review of Research on the Neurocognition for Timbre Perception[J]. Frontiers in Psychology, 2022(13): pp1-11.

第二章 人类的语声

本章从人体生理结构入手，对人类发声的构造和原理进行简要概述，并延伸至对人声与体型对应关系的分析。通过本章，读者能够更清晰地理解语声是如何产生的，语声各要素能够传达何种信息，这为后续实验提供了科学的理论基础。

第一节 人体发声机理

人类语声的产生可以被视为一个四阶段的过程（图 2.1.1）：

- (1) 呼吸系统压缩空气，产生气流；
- (2) 震动的声带将该气流转化为间歇性或脉冲性的经喉气流，生成一个复合音，即声源；
- (3) 共鸣系统对此复合音进行滤波；
- (4) 大脑支配口腔内不同部位协作，将声音转化为语声。

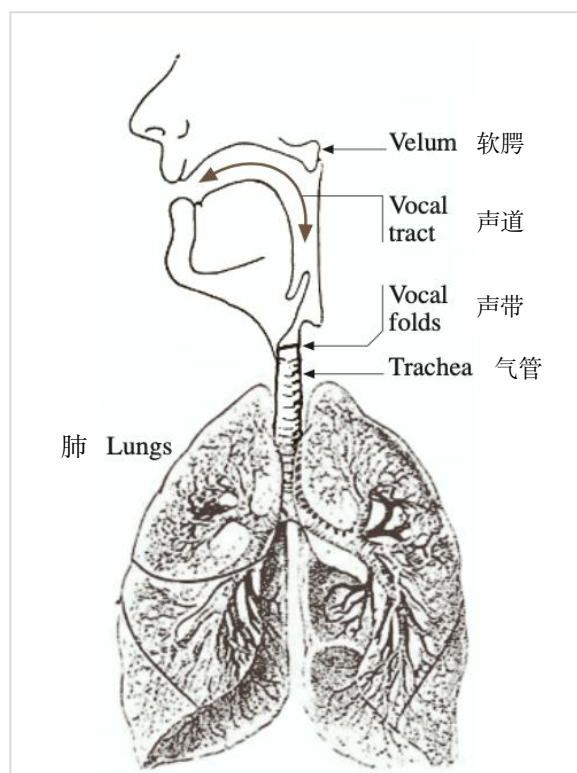


图 2.1.1 人类发声系统结构图^①

^① M. Schroeder, Thomas Rossing, F. Dunn. Springer Handbook of Acoustics || The Human Voice in Speech and Singing[M]. NY: Springer New York, 2007: pp669-712.

一. 呼吸系统——能量

呼吸系统为发声提供所需的气息能量。人的呼吸系统由鼻腔、口腔、喉部、肺部以及横膈膜相关肌肉组成。其中，肺部是人体发声系统的动力源^①。连接肺部和喉部的是输送气流的气管，气流最终通过气管、支气管到达喉头，为声带调节提供能量。

二. 声带震动——声源

呼吸系统传送至喉头的气流会引发声带震动，从而形成声源。在这一过程中，声带的长短、薄厚、松紧等一系列形态特征都会影响声源的声学特征。同时，稳定气流在声门（Glottis）处会因声带振动而被转化为周期性的脉动气流^②，喉部通过此气流控制声门的开合形态，这一结构也会对声源参数产生影响。

三. 共鸣系统——滤波

声道包括口腔、鼻腔、咽腔三个部分，属人体共鸣系统的上部共鸣体。声源从声带、口腔后部的软腭、到牙齿之间的流转需要相当长的时间，过程中声音所经各部位都会对其频谱产生不同程度的影响，这就是著名的“源-滤波器模型”

（1960, Fant）。人体声道比稳定不变的物体复杂得多，黏膜的状态、弹性、能够支撑声道的肌肉、软骨劲度，每一时刻都在发生细微的变化。实际上，可以将声道理解为一个物理系统，它的物理特性决定了人声的共振峰参数特性，对最终音色的形成至关重要。

人体的共鸣系统除声道外，还包括下部共鸣体——胸腔，这一共鸣体常常作用于低频。当我们降低音调讲话时，能够感受到强烈的胸腔振动，这是由于受到支气管压迫的空气在胸部形成支点，从而引起的胸腔共鸣。反之，当我们提高音调，甚至使用假声说话时，随着音调的升高，胸腔的共振会越来越弱，此时上部共鸣体的振动也会逐渐增强^③。

从生理角度看，音色的形成除了受声带长短、宽窄、厚薄和闭合状态影响外，更重要的是与共鸣腔体的物理形态、调节程度密切相关^④。共鸣系统能够反映人体发声的整体共振特性，影响原有的频率分布，是音色形成过程中的关键一环。

^① 李晓辉. 汉语语音特点的分析及应用[D]. 北京邮电大学, 2014.

^② 陈小平. 声音与人耳听觉[M]. 北京: 中国广播电视出版社, 2006.

^③ 张悦. 计算机环境下人声合成技术的艺术化运用[D]. 南京艺术学院, 2016.

^④ 王冬梅. 论声部划分的依据[J]. 才智, 2011(10): 1.

四. 构音系统——语言

构音系统用于塑造发音部位，产生可供识别的语声。人类之所以能够发出不同的元音和辅音，主要依赖于口腔的运用，它不仅是共鸣器官，也是发音器官。当我们说话时，通过中枢神经支配，口腔内的上下唇、齿、舌头、软腭等部位彼此协作^①，各组织形状、张力、位置随之调节，使得共振峰在频率上发生改变，从而形成我们想要的音节，并进一步组成语言。

第二节 人声频谱与体型的对应关系

在大多数陆地哺乳动物中，发声的基频 (F_0) 和共振峰间距 (ΔF) 与生物的体型大小呈负相关^②，较大的生物体一般倾向于比较小的生物体产生频率成分更低的声音。但这并不绝对，只是一个总体的趋势。

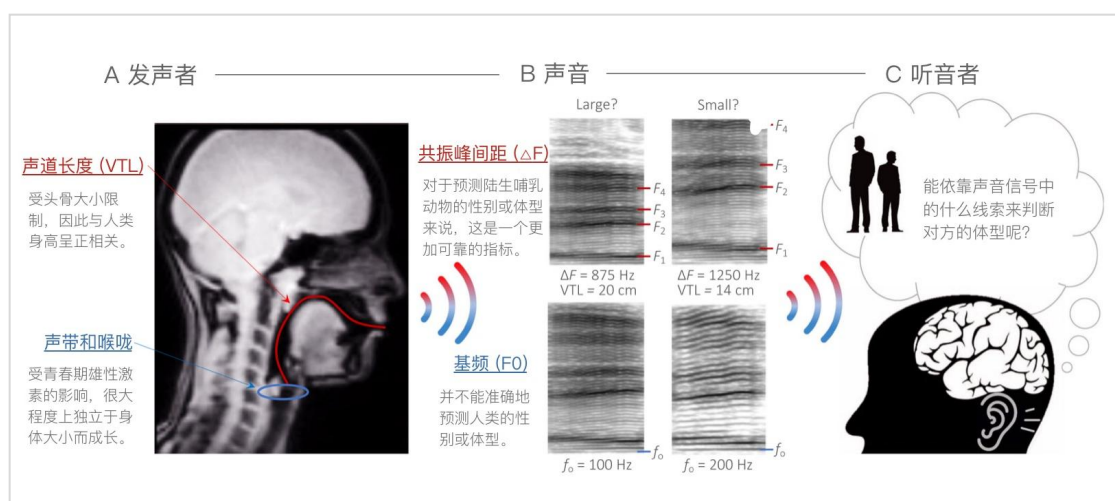


图 2.2.1 与体型相关的声音信号传递链^③。

A 为人类声带解剖的矢状核磁共振图像，显示发元音/u:/时的状态；
B 为体型大、小两人发出同一声音时，各自的基频和共振峰（标记为 F_1 - F_4 ）。

人类的声道位于头部范围内，颅骨大小和身体大小紧密相连，因此体型更大的人通常会伴有更长的声道，这导致了声音的共振峰频率更低，共振峰间隔也更

^① F.Alton Everest, Ken C. Pohlmann. 声学手册（第 5 版）声学设计与建筑声学实用指南[M]. 北京：人民邮电出版社, 2016.

^② David Pietraszewski, Annie E. Wertz, Gregory A. Bryant. Three-month-old human infants use vocal cues of body size[J]. Proceedings of the Royal Society B: Biological Sciences, 2017, 284: 20170656.

^③ Katarzyna Pisanski, David Reby. Efficacy in deceptive vocal exaggeration of human body size[J]. Nature Communications, 2021, 12: 96812(1).

为密集^①。在如考拉等一众哺乳动物的相关研究中，动物叫声的共振峰参数也与体型大小表现出了显著的关联性^②，同样证明了共振峰信息可以成为判断发声体体型信息的重要指标。

人声的基频，理论上是指由声带震动产生的频率，它对音高的决定作用使得大众普遍认为，那些听起来低沉、厚实的声音直接指向了更低的基频。但事实上，这一指征在众多动物解剖学研究中都被证明并不能与发声者体型大小准确对应。这是因为产生基频的声带由声带肌、声带韧带和声带粘膜三部分组成，这部分软组织的发育程度仅受雄性激素影响，独立于骨骼结构，故不能作为判断身高和体型的可靠线索^③。但从人类两性差异的研究来看，发育后男性的声带普遍比女性长，因此基频也会普遍更低^④，同时在长久以来的视觉经验中，男性的体型相较于女性普遍又更加高大，这种视觉和听觉的稳定关联可能会说服我们不自主地将低基频与男性化特征联系在一起，从而导致了这种基频与体型的错误对应。总的来说，低基频可能会使发声者听起来高大、强壮、具有攻击性，但它与身体大小的关联程度要远弱于共振峰参数。

综上所述，体型较大的人类往往匹配的是较大的声压级、较低的共振峰频率、较密集的共振峰分布，以及可能较低的基频。此理论阐明了人类体型特征与声学特征之间的相关性，这二者的对应关系与本文论题息息相关，应被用作语声创作的重要依据。

^① T. Fitch. Encyclopedia of Language & Linguistics (2nd Edition)[M]. Amsterdam: Elsevier Science, Editor(s): Keith Brown, 2006: pp115-121.

^② Benjamin D. Charlton, William A. H. Ellis, Allan J. McKinnon. Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: Honesty in an exaggerated trait[J]. Journal of Experimental Biology, 2011, 214: pp3414-3422.

^③ Stephan Reber, Judith Janisch, Kevin Louis Torregrosa. Formants provide honest acoustic cues to body size in American alligators[J]. Scientific Reports, 2017, 7: 1816.

^④ Yi Xu, Albert Lee, Wing Li Wu. Human Vocal Attractiveness as Signaled by Body Size Projection[J]. PLoS ONE, 2013, 8(4): e62397.

第三章 《沙丘》中哈克南男爵的对白复现实验

基于前文对于“类人”语声特征和人类发声理论的相关研究，“类人”角色语声形象的创作方向已经初具雏形。为了更加深入地探究这一问题，本章将以科幻电影《沙丘》（*Dune*, 2021）中的哈克南男爵一角，作为“类人”角色中对人类声音特征保留得较为极致的特例展开讨论。旨在通过以理论为依据的推测、相关创作者及同类创作经验的归纳整理，以及一系列具体实验，最终提炼出一些行之有效的创作方法，以便声音创作者们能够在理性分析的基础上，科学选用处理手段，做到思路明晰、有的放矢，从而创造出写实性与表现力并举的“类人”语声形象。

第一节 复现实验思路依据

由于《沙丘》声音创作的精彩之处甚多，声音团队尚未在任何公开访谈中提及哈克南男爵语声的创作，况且影片风格不同，其处理方式也往往不尽相同。因此，为了更加全面、准确地研究此类型声音的创作，笔者尝试在可实现的方式中尽可能拓展更加多元化的思路和方法，以下将从三个角度逐步展开。

一. 基于角色背景和视觉形象的语声构思



图 3.1.1 漂浮的哈克南男爵

漂浮在半空的弗拉基米尔·哈克南男爵，是哈克南家族的统治者（图 3.1.1）。他拥有传奇版的邪恶和智慧，被称为“所有科幻小说中最具标志性的恶棍之一”。残忍、狡猾、邪恶、贪婪是哈克南男爵的核心特征，他极度擅长操纵他人弱点，对服务于家族的奴隶实行暴虐的统治，判处年轻的保罗死刑，甚至牺牲自己的侄子，与此同时，他也具有相当的智慧和野心，通过运作香料流动，为自己积累了巨大的财富和权力^①。

关于男爵的原始外型设定，弗兰克·赫伯特（Frank Herbert）在《沙丘》的同名原著中写道，“病毒在今后的数年内把弗拉基米尔英俊的外表变成丑陋、肥胖、满身脓疮的样子，最后他的双腿甚至无法支撑起巨硕的身体重量，以至于需要一种反重力浮空装置来减轻行动的负担。”原著中的文字描述为读者刻画了一个生动、具象的画像，也为男爵形象的塑造提供了基础。

画面中的视觉形象是角色呈现的核心，因此外形的实现方式往往也可以被视为语声设计的重要线索。从男爵特效化妆的幕后访谈^②中能够得知，画面中怪兽般的体型没有使用任何数字特效，是完全依靠手工填充硅胶肢体完成的，演员的脸上藏有七块硅胶，仅指关节上的手部硅胶就重近 10 磅。之所以选择这样费力的塑造方式，是因为导演维伦纽夫不想让男爵看起来像一幅漫画，“他要充满整个房间，我想真正感受那个巨大的人类和那个巨大的重量带来的威胁^③。”这些关于美术造型的创作思路也为语声形象的设计提供了参考。

综合以上有关故事背景、形象设定、外貌体征等信息，哈克南男爵完全符合本文第一章对于“类人”角色的定义，其语声创作也应遵循前述对“类人”角色语声特征的分析。由此，笔者得出哈克南男爵语声形象的设计目标：庞大、深沉、威严，与人声相似又相异，充满黑暗魅力。他要表现出一种你从未见过的存在感，而不仅仅是像一个怪物。

二. 对于原声带的分析

原片的声轨是创作结果最直接的呈现，也是实验重要的研究依据。通过对声轨^④不同方面的剖析，往往可以收获颇有价值的信息，以便进一步推测其创作思路。以下将从声道内容和语声信息两方面展开。

^① Baron Vladimir Harkonnen Vladimir Harkonnen[EB/OL].
https://villains.fandom.com/wiki/Baron_Vladimir_Harkonnen.

^② Bryan Alexander. The making of oversized ‘Dune’ villain Harkonnen: no CGI, just a lot of prosthetics[N].
Chicago Sun-Times, 2021.10.29.

^③ Marcus Gabriel. Dune Behind the Scenes: Creating Baron Harkonnen[EB/OL]. 2021.9.30,
<https://dunenewsnet.com/2021/09/dune-behind-the-scenes-creating-baron-harkonnen/>.

^④ 注：分析对象为 7.1 声道版本文件。

1. 声道内容

从声像分配来看，男爵的语声信号以 C 声道为核心，LssRss 声道次之，LR 声道则并不作为延时声发送的主要对象，总体呈现出一种“横向延展”的形态。

以实验段落中“House Harkonnen would never dream of violating the sanctity of your order.”一句为例进行分析，选用插件 Insight，分析界面为电平表和声场图，设置峰值冻结模式。由界面数值可知（如图 3.1.2），C 声道峰值约为-10 dB，LR 声道在包含持续性音乐的情况下仍然仅-33 dB，而 LssRss 声道则高达-12 dB，且音色干净结实，混响尾部衰减几乎不可闻，在时间差方面较 C 声道延迟了约 40 ms。因此，推测是设置了单次激发的短延时。



图 3.1.2 Insight 分析界面

此外，在整个段落中，每当角色发出声音，LR 声道和 LsrRsr 声道就会出现与之同步的 Buzz 噪声，并且仅与对白声同步出现，效果声并不引发同步的 Buzz。该噪声频率范围约为 5 kHz - 20 kHz，频响基本平直，音色与小军鼓滚奏相似。推测是用 C 声道的声音内容侧链激发了某种噪声信号，但其使用目的不详。

2. 语声信息

通过不同段落的对白表现可以感受到，男爵发声时，声带厚实、松弛，规律性振动的频率较低，音色结实有力，尤其在发出元音音节时，胸腔共鸣浑厚，鼻腔共鸣稍弱，声带振动和声道共鸣的特征表现更加明显，会产生带有沙砾感的低沉颤音。男爵语声的整体听感完全符合人声特征，但体型会远大于一般人类。

从语声特征的其他方面而言，男爵说话时多用陈述句，语速平缓稳定，音节

持续时间长，语调起伏较小，重音清晰明确，这都辅助了对于男爵沉稳威严的个性、运筹帷幄的心态和极高社会地位的塑造，呈现出强烈的威胁感和掌控力。

基于第二章对于人类声音特征和发声原理对应关系的分析，并结合男爵语声中实际传递出的信息，笔者推测，男爵语声形象的塑造可能是：在录制时通过非常规的话筒类型或拾音位置来强化原声腔体的震动感；后期在频率上对基频、共振峰、谐波进行细致调整；在动态上进行压缩强化力量感；添加延时混响或激励次低频表现角色体型；添加失真赋予语声更强烈的粗糙感和情绪表现等方式来实现的。

三. 参考混录师 Ron Bartlett 所述

通过对角色背景下语声特征的构思以及对原声带的分析，我们能够初步推测出拾音方式、音色塑造、声道分配等方面大致选用的手段，但这并不足以明确对于角色语声的具体处理思路，而这恰恰是最关键的部分。

因此，笔者尝试联系到《沙丘》的声音团队，希望能更为清晰地了解创作者在动作背后的考量。在后续与混录师 Ron Bartlett 往来交流的邮件中，Ron 慷慨地介绍了他对此声音的处理思路、手段及看法（详见附录 1）。

简要概括为以下 4 点：

1. 以激进的 **EQ** 设置，增强某些音调，扩大声音范围。
2. 以准确的**压缩器**设置，挑出原声的特征音色。
3. 以特殊设计的**延时**，增大角色对于空间的影响。
4. 以编辑音频**增益**，强化角色的语气和压迫感。

邮件所述的前三点均于笔者前文分析契合，第四点则提示了对特定音节进行编辑的重要性，进一步细化了处理精度。与前述推测相异的是，Ron 并没有使用非常规的拾音方式，也没有添加失真和次低频来辅助塑造角色的语声特征。他在个别音节上的确使用了变速变调，但处理目的仅在于匹配口型。此外，Ron 还特别提到，低频增强声道的使用会让对白变得糟糕，这提醒了我们必须时刻注意掌握表现力和清晰度之间的平衡，在加深、增厚男爵语声之前，应首要保持可懂度和人类特征，尽可能选用更加细腻的方式来挖掘对白中的细节。

考虑到不同发音者有着不同的生理特质，其差异性必然会导致同样的手段并不一定能得到等同的处理效果。因此，关于特殊拾音方式、变速变调、失真等方法，仍旧保留着应用的可能性，实际效果有待结合实验进一步验证。

第二节 拟定实验方案

一. 录音阶段

1. 拾音方式

主话筒选用同期常规话筒，布于演员头部前上方，与地面夹角约 45 度指向演员嘴部，拾取清晰、自然实验基础素材。同时，可根据需要选用特殊话筒配合主话筒同步录制，为后期处理提供更丰富的音色选择。笔者在综合参考人体发声原理、话筒拾音特性及相关经验分享后，拟选用接触式话筒作为补充。接触式话筒 (Contact microphone)，一种通过与固体接触来感知振动的特殊话筒。独特的拾音原理让它能够捕捉到一些非声波式的、物体内部的微妙声音，这也是声音设计师们多年来一直将其视为重要创作工具的原因。依据前文对于人体发声机理的研究，胸腔作为下部共鸣体，其震动对于低频声音的塑造有着很大的影响。为模仿这一特征，实验拟将接触式话筒粘贴在演员的喉部或胸部，与主话筒同步录制，拾取身体因发声而产生的真实共振细节，为塑造体型和音色特征奠定基础。

2. 高采样率录制

在设备条件允许情况下，考虑使用高于 48 kHz 的采样率进行录制。这样在后期无论是需要直接调整音高，还是由于匹配口型的需要，对部分字节进行变速变调处理，都可以更好地保持原声的音色和细节，降低失真。

二. 后期处理阶段

1. 音色表现

此步骤的目的在于为男爵塑造特殊的“类人”音色，通过对原始素材进行一定程度的变形处理，使之成为有别于人声的“类人”语声。拟采用**压缩**、**均衡**两种手段，以更加极端的参数设置来强化语声中的特征音色，甚至赋予新的音色。

2. 体型表现

此步骤以人类发声机理、人声特征和体型特征的对应关系为依据，在保留基本声音特征的基础上，尽可能夸大男爵语声的体积感和重量感，传递男爵最具代表性的外形信息，以语声形象的表现力为视觉形象增色。

拟采用以下手段：**第一，形成共振峰**。体型更大的人通常伴有更长的声道，会导致声音的共振峰频率更低，共振峰间隔也更为密集，这是前文的重要结论。

基于这一理论，实验中拟使用均衡器对素材的谐波信号进行一定程度的调整，通过塑造新的共振峰分布形态，辅助表现大体型发声者所对应的声学特征。**第二，独立处理低频。**声音设计师 Tzvi Sherman 曾在一个关于巨型外星蠕虫的声音创作分享中^①提到了这种增厚低频的方法。为进一步增强声音传递的重量感，拟将素材复制一份单独放置在新的轨道上，截取低频部分进行一系列强化处理，最终再与原声叠加。**第三，创造人体内部的延时混响。**由第二章理论分析结果可知，更庞大的体型并不一定会产生更低的基频，但一定会携带更宽阔的共鸣腔体。因此若能模拟出声带振动后，在更大空间反射共振产生的厚实共鸣，对于表现夸张化的角色体型必定有所帮助。拟将大空间混响、小空间混响、短延时等多种信号叠合，以此来模仿胸腔参与共振后产生的巨大低音。**第四，降低音调。**虽然相关理论已经证实了基频与体型之间的关联实际并无必然性，但考虑到变速变调是声音制作中惯常使用的手段，在动物声、效果声等各类声音元素中都有着广泛的应用，因此这一方式依然留有实验价值。拟通过降低音调来变形语声，检验这一处理是否能够辅助语声的体型表现。

3. 情绪信息表现

此步骤旨在赋予男爵语声更多的感性特征，辅助传递角色的社会地位、身份背景、性情心智等潜在信息。拟采用以下两种手段。**第一，编辑增益。**选择素材中需要加重情绪的字节或音节进行音量上的细微增益，这不仅有助于对白语气的强化，也能够更好地表现处理后的音色特征。**第二，添加失真。**这一种更常用于效果声的处理手段，能够很好地增添声音的冲击力。拟将较低量值的失真信号应用于对白声音，丰富男爵语声中关于统治力、攻击性等情感特征的表现，实际选用的失真类型、处理方式和最终效果有待实验进一步检验。

第三节 实验过程

实验音频来源：一位嗓音偏低，英文口语好的男性同学。

实验音频内容：该同学对片中哈克南男爵出场段落对白的模仿配音，共六句，中英双语。

以前述思路为依据，笔者在归纳、整理，并结合实际情况适当调整后，对中英文录音展开了如下方式的实验。

^① Tzvi Sherman. Alien Worm Creature Vocals Sound Design[EB/OL]. 2021.1.15, <https://www.youtube.com/watch?v=1qlO7LTF0qk>.

(1) 叠加水听器 (Hydrophone) 录音

因实验条件所限, 此步骤无法依照预期方案使用接触式话筒进行辅助录音, 笔者在仔细考察其它各类特殊话筒的拾音特性后, 最终选定了与接触式话筒原理相近的水听器进行实验。水听器虽然专为水下使用而设计, 但那些在水环境中有效工作的特性, 使得它也可以移步换形, 应用于离水的气体环境^①。

实验中, 将水听器湿端紧贴演员喉部进行录制。

通过实验, 笔者意识到这样的使用方式存在两个问题。第一, 非水环境中, 水听器表面与声源无法完全耦合。对于湿端而言, 实际有效的拾音区域远小于设备所供, 大部分只能暴露于空气中, 因此无论是接触面积的限制, 还是空气粒子对剩余透声部分的持续碰撞, 都会影响最终的信噪比。第二, 水听器所使用的压电陶瓷传感材料是针对液体环境设计的^②, 与其它介质的声阻抗匹配度很低, 这也会对声音传输产生一定干扰。但即便如此, 声音所呈现出的质感和音色依然有着普通话筒无法实现的特殊性, 为男爵增添了很强的异化特征和攻击性。

(2) 降低音调

在听感可接受范围内, 降调的极限数值仅 15 个音分左右, 处理后音调能够降低, 但基本不对体型表现产生作用, 故不保留。

降调过程中, 声音的基频和共振峰会一同按比例降低, 这是由变调效果器的工作原理导致的。这样的特性对效果等一类声音的使用并无干扰, 甚至可以化劣势为优势为声音增加冲击力, 但语声信号的复杂性和多变性要远高于其他声音, 发声者每一时刻的喉颈状态、软腭高度、舌位、口型变化都会形成完全不同的共振峰形态^③, 同一元音又有着固定对应、不随音调变化的共振峰。如此复杂的人体规律, 使得处理稍有不慎就会造成明显的“怪异感”。仔细听音调降低的过程, 可以感受到发声者口腔肌肉的松弛程度也在逐渐减弱。

(3) 处理共振峰

借助均衡器调整共振峰形态。首先使用较大 Q 值提升或衰减相应频段, 塑造更密集的共振峰分布, 同时提升低次谐波, 避免低频能量聚积在基频附近。在这一步中, Fabfilter Q3 的实时频谱功能可以大大增加处理的准确性和效率。

处理后, 语声在保持清晰度的同时, 传递出了微妙的体型信息。

^① Brüel & Kjær. Hydrophone[EB/OL].
<https://www.bksv.com/zh/knowledge/blog/sound/gas-environment-hydrophone>.

^② John A. Maurer iv. Research in Underwater Sound:With Focus on Musical Applications and Computer Synthesis[EB/OL]. 1998.3.15, <https://ccrma.stanford.edu/~blackrse/h2o.html#hydrophone>.

^③ 杜士武. 咽喉机能原理与嗓音声学分析[M]. 吉林: 吉林大学出版社, 2020.

(4) 均衡

首先使用截止频率为 80 Hz 的高通滤波削弱原始素材中的噪声，为后续添加低频谐波成分留有余量。随后，找到中低频段共鸣最强的频点进行增益，并尝试对其倍数谐波成分进行调整，程度稍弱。此步骤对精确度没有过高要求，可考虑选用 Scheps 73、PuigTec EQP-1A 等模拟均衡器，对声音质感的提升效果更好。

处理后，声音的温暖感和共振感明显提升。

(5) 平行压缩

将原音频复制一份放置在新轨道上，对复制音频进行较为激进的压缩处理，找到能够强化特征音色和瞬态细节的参数，再将处理后结实肥厚的压缩信号与带有音头瞬态的原信号叠加，最后整体调整两信号的比例关系。

需注意是否出现抽吸现象。实验中，经过前述处理对中低频的提升，压缩时产生了明显的抽吸感，后通过侧链功能提前将部分低频滤波得以改善。

处理后，声音的色彩感明显提升，音腹也更加丰满。在这一步中，可依需选择不同类型的压缩器以获得不同的压缩质感。笔者在比对后选用场效应管压缩器 CLA-76，以较慢的启动时间让出音头，中等释放时间保留体积感，压缩比为 12:1。

(6) 独立处理低频

将原音频复制一份单独放置，并对其进行 250Hz 的低通滤波，仅保留低频段待处理。随后将 120 Hz - 180 Hz 频带范围略微提升，再将其送入次低频激励器中，激发出更为饱满、松弛的次低频信号。与原声叠加后，的确获得了比全频带处理更好的体型塑造效果。

(7) 添加用于体型塑造和空间辐射的特殊混响

对于体型塑造而言，混响声应以反映人体内部空间为目标。使用原音频激发单声道小房间混响后送入 C 声道，混响辅助轨道需设置 1 kHz 的低通滤波预处理，并添加噪声门使混响信号在干声结束的同时迅速截断。对于空间辐射而言，笔者参考前文对于原声带的分析，激发 Ambience 类型的混响后分别送入 LR 声道和 LsRs 声道，后者的送出量值要大于前者。

混响信号的加入，极大辅助了角色的体型表现。

(8) 编辑增益

使用 Protools 的音频增益功能，挑选多处可以更好表现情绪以及音色特征的音节，略微提升电平，如第一句中的“die”、“atreides”的音头部分。

处理后，语气中的情绪表现更加明显，前续步骤塑造的体积感、重量感、共振感也获得了更进一步的优化。

(9) 失真

使用 Abbey Road Saturator 饱和效果器，预处理 EQ 设置截止频率为 2 kHz 的低通滤波强化中低频，饱和参数设置为 30，饱和风格选择音色更加圆润的 TG12345，最终与原声调整混合比例。

添加轻微的饱和失真后，语声的暖度和力度明显增加，这很好地强化了攻击性和体型表现，饱和所带来的高频细节，也能在一定程度上补充前述处理造成的高频损失。此步骤需特别注意尺度，因为添加过多饱和反而会弱化瞬态，降低清晰度和冲击力，甚至让声音变得“嘈杂”，具有咆哮感。如果需要更轻微的效果，可以选择让信号经过模拟效果器来为声音增加一些谐波成分并控制 THD 量值，例如结合了 1066 和 1073 效果器的 VEQ-3。添加轻微的谐波失真会让声音更有光泽感。这种处理较为细腻，对声音的劣化程度低于饱和失真。

第四节 实验结果及分析

经过对上述步骤的多次调整和尝试，实验在塑造男爵音色特征和体型特征方面取得明显效果，基本符合角色表现所需，但是对于混响部分的处理，仍未能很好地反映出角色发声对于整个空间的影响，语声的清晰度和颗粒度也与原片存在差距，尚待进一步研究。整体来看，英语版本的处理效果要好于汉语版本。

第四章 拓展使用针对性复合技术手段

在复现实验中，笔者以不同处理阶段、处理目的为依据，对人声各个特性的处理展开了较为详尽的分析，以便读者能够更清晰地了解每一步骤的细节及原理。但与之对应的是，实际创作过程中，如此庞杂的处理流程和零散的插件使用往往带来的是更低的工作效率、更高的沟通成本以及更大的学习难度。因此，在掌握各类声处理原理和使用规律的基础上，我们可以将目光转向更为高效统一的、针对性更强的复合效果器。

针对语音处理的复合效果器中融合了多种技术手段，能够从不同的技术角度和创作思路对人体发声器官进行针对性的模拟和变形处理，还能够在此基础上对不同生理结构所对应的不同声学特征展开相对自由的调整。此外，不同的复合效果器会依据其目标的应用场景和处理方向，对内部集成的处理模块嵌入不同特色的优化设计，这使其相较于前文所述的独立手段具有更强的适用性。

本章将拓展介绍 VocalSynth、Throat、Dehumaniser、OVox 四款针对语音处理开发的复合效果器。通过对“类人”语声两种特征的进一步实验，总结每款复合效果器的处理效果和对应处理思路，并尝试对其风格和适用性方面进行初步的对比和归纳，希望能够为实际创作中实现快速、准确的效果器选择提供支持。

第一节 两种特征的实现效果对比

一. 关于“类人”音色特征

不同的效果器拥有着对声音进行不同方式、不同风格、不同程度变形的能力，这极大地影响了“类人”音色特征的塑造方向。在对比实验的过程中，四款效果器都表现出了其各自的处理特点，其中，VocalSynth 和 Dehumaniser 对于男爵音色特征的实现效果更加理想。

VocalSynth 的五种发声工具都能够高质量地改变音色特征，其中，Biovox、Compuvox、Polyvox 的表现更加契合本文实验的处理需求。以 Compuvox 为例，这一模块基于 LPC 技术，能够将声音信号分解为共振峰和激发源两个部分独立运算，所提供的拉长元音持续时间的 Bytes 参数、赋予蝙蝠侠般沙粒感的 Bats 参数，也都可以为声音添加独特的细节变化，效果极佳。但通过谐振滤波器加入元音共振峰的设想并未实现，这种方式对语声自然度产生了较大干扰。

Dehumaniser 的处理风格更加强烈，但把握尺度仍能达到理想的效果。实验选用 Dialogue 类型中的 Cracked up 预设，通过两个采样触发器调节音高和延时，通过噪声发生器发出中心频率为 980 Hz、带宽为 8000、谐振为 0.1 的噪声信号，并将这三个模块并联排列于输入信号后端，最后再调整与原声之间的比例关系。这种方式有效地加强了男爵音色的侵略性，甚至为其赋予了一些微妙的病理信息。需要注意的是，Dehumaniser 追求对声音的彻底改变，若创作目标为与本文实验目标相似的效果，应谨慎选择环形调制器、声码器等对自然度破坏较大的处理手段。

而 Throat 的设计目标则仅针对声带塑形，因此在音色塑造的多元性上同样有所局限，并且当参数设置较大时对声音的劣化严重；OVox 的处理风格则偏向于人工化、音调化，在创造自然“类人”音色的实验中效果欠佳。



图 4.1.1 VocalSynth 界面



图 4.1.2 Dehumaniser 界面



图 4.1.3 Throat 界面



图 4.1.4 OVox 界面

二. 关于体型特征

体型特征塑造的难点在于精确性，这要求它的处理必须与人类生理构造紧密关联。实验发现，基于人类声道物理建模的 VocalSynth 和 Throat，对于这一特征的呈现效果更佳。

VocalSynth 的可调参数多，处理细腻，其参数设计跳过了生理指标向声音特征转化的过程，更加关注对声音的主观感受。实验通过对 Biovox 模块和 Polyvox 模块的联动调整，实现了较为理想的结果。其中，Biovox 模块的 Nasal 可以增加鼻音，有助于模拟软腭形态的改变；Breath 可以增加高频信息，有助于塑造近距离的压迫感；但 Shift 对体型的表现不如 Polyvox 模块中的 Formant 自然。

使用 Throat 处理时，拖拽节点直接对应器官形态的功能大大强化了处理的有效性，尤其是在角色生理具有明确特征时。实验中，笔者通过调整五个发声部位的长短、宽窄，并结合嗓音范围、声门模型、喉部精确度、脉冲带宽四个参数，大体达到预期目标。Throat 的独特设计为我们提供了直观、便捷的方式，也能够一定程度上辅助表现体型特征，但其对声音细节的保留并不理想。

Dehumaniser 和 OVox 虽然也能够塑造体积感和重量感，但它们的调整更为激进。例如如 Dehumaniser 仅能通过组织声码器、变调器等对声音改变较大的模块来塑造体型，处理精度不足，也没有与生理特征相关的参数可供调整，此特征细腻、自然的处理目标并不契合这类效果器的特性，其最大的优势还是在于对声音进行彻底变形和驱动声音之间融合的能力。

第二节 处理优势与适用方向

上述两种特征实现效果的差异，很大程度上是由效果器的设计思路决定的。对于同类型的处理手段而言，其功能布局、可调参数、信号路由方式、不同模块间合作机制的区别，实际上体现的是设计思路的区别，这能够准确地反映设计者对于所开发效果器应用目标的倾向性。

VocalSynth 最大的特点在于它的处理精度。5 种不同类型的发声工具，以及为单一类型所设置的众多细节参数，使得 VocalSynth 能够基于发声原理、生理结构来对不同声音进行变形，这种设计思路精准、自然，在用于人声处理时具有显著优势。

Throat 在设计重心上与 VocalSynth 有着一定的相似性，但其更加擅长于对发声结构进行直观调整。它允许用户直接选择声音风格、声门闭合程度等预设，并支持手动拖拽节点来改变对应发声器官的长短、宽度、形状，效果器会实时运算出器官形态所对应的声学特征。总的来说，Throat 与生理参数的关联性较强，对人声的调整较为科学，更适用于塑造与发声器官生理状态密切相关的特征，如性别、年龄、病理情况等。

对于 **Dehumaniser** 来说，高度的创作自由是构成其独特性和复杂性的核心，这源于它“自主定义”的设计逻辑。Dehumaniser 的本质是一个可供自由调配的

“节点系统”，用户可以完全控制信号流的路径，可以同时使用多个模块，甚至设计复杂的串联或并联结构，最终创建出一个自定义的动态互连系统。除此之外，Dehumaniser 用于创作还有一个显著的优势在于，它能够对话筒输入信号进行实时处理。比起手动调节素材库中的内容，用自己的声音直接驱动会更加灵活，这不仅是因为过程中可以依据听感及时调整发声方式和力度，也在于人类声音的介入会让最终的声音更加富于情感和变化，在赋予兽化特征的同时，保留语言所携带的语调、语气、情绪等信息。《怪奇物语》中，“夺心魔”的对白就是 Craig Henighan 以自己的声音驱动处理后，再进一步融合狮子、老虎、海象和一些干冰的声音创作而成的^①。与前两款复合效果器相比，虽然设计目标都是对人声的变形、异化，但 Dehumaniser 独特的信号路由方式、预制类型和处理特性，都使得它尤其适合于制作生物、怪兽、机器人等夸张化的声音。

OVox 的界面设置和处理流程都和 VocalSynth 有着很高的相似度，但与 VocalSynth 不同的是，OVox 更加看重合成技术的使用，其工作原理也是基于一种有机再合成 (ORS) 技术，通过将原始信号分解为 3 个核心信息：振幅、音高、共振峰，再重新合成来创造新的声音^②。此外，OVox 还提供了融变 (Morph)、调谐、和声，以及 Talkbox 等一系列偏离自然效果的处理手段。这些独特的内部设计使得 OVox 更适合处理音乐、机器人等一类对合成感、电子感容许度较高的声音。

针对性复合效果器的应用大大降低了手动建立复杂处理链路的难度，这有助于我们快速进入创作性的工作。但很多情况下，为了增加角色声音的个性和表现力，我们往往寻求的是一种复杂、糅合的声音，这时候就需要综合使用各类技术手段，依据其各自的处理优势进行组合，并参考第三章思路，对某些特征进行针对性的补充处理，最终创造出一个独特、细腻、契合角色需要的声音。

^① Jennifer Walden. Behind the Growing, Evolving Sound of ‘Stranger Thing’ Season 3 – with Craig Henighan, Will Files and Mark Paterson[EB/OL]. 2020.7.10, <https://www.asoundeffect.com/stranger-things-season-3-sound/>.

^② Waves Audio LTD. OVox Vocal ReSynthesis User Guide[EB/OL]. 2020, <https://www.waves.com/1lib/pdf/plugins/ovox-vocal-resynthesis-v2.pdf>.

第五章 思考与应用

前述章节通过细致的理论分析和具体实验,得出了以人类生理特征与语声特征的对应关系为理论依据,能够有效辅助“类人”语声创作的结论。本章将在此基础上,从应用于实践过程、应用于中文语声、应用于其它声音三个方面,进一步讨论这一结论是如何联结实际创作的。

第一节 实践中的应用要点

具体处理前,首先要依据角色外型、故事背景、影片风格等方面对目标语声进行初步构思,考虑角色的语声形象应大体反映出其外型形象的哪些特征,并依据此拟定录音方案。其次,在录音过程中,应参考人体发声机理,对多种话筒类型、拾音位置、拾音方式展开广泛尝试,为后期处理提供更丰富有效的素材选择。最后,应对录制成果进行仔细地聆听、判断,辨别原素材与目标语声之间的差距,再结合生理特征与语声特征对应关系及实践经验,筛选出合适的声音素材。

后期处理过程中,应在对角色定位和发声原理进行理性分析的基础上,依据创作经验和具体判断,选择性地混用各种处理手段,并尝试不同处理顺序带来的效果差异。同时,务必注意的是效果器对声音素材可能产生的副作用,少即是多,一切处理以对白的清晰度和音色特征的保留为首。

电影声音是综合艺术,需要各类声音元素之间相互配合才能形成一个协调的整体。从这一思路出发,除了上述众多方式能够对语声产生实际变形外,通过同一场戏不同角色间的音色对比、响度对比等间接手段,也能够有效地强化听感,还可以降低对目标语声处理强度的要求,为更保守的参数选择留出余地,进而更好地保护语声信号的细节。

第二节 汉语与英语的应用差异

事实上,不同语种的母语者在发声时具有潜在的生理差异,这导致本文以英语为例的实验所提炼出的经验会存在一定的局限性,无法精准适用于其它语种。考虑到这一问题,本节旨在原有结论的基础上进一步对汉、英两语种的发声特征展开分析,并初步讨论有助于补充语声特征差异,从而让汉语能够更好地应用本文所述思路的处理手段,希望能够为中国电影的语声创作提供参考。

一. 关于语言特征

1. **语言系统的根本差异。**汉语是声调语言 (Tone language), 声调在汉语中作为音位存在, 具备区分语义的功能。比如“衣” yī、“移” yí、“椅” yǐ、“义” yì 四个字, 音位完全相同, 仅凭声调差异就构成了完全不同的字义。而英语是语调语言 (Intonation language), 声调在英语中仅负责传达语气, 其本身并不具有含义, 这一点是汉、英两个语言系统最主要的区别^①。在此基础上, 两语种逐渐演化出了不同的语言特征, 汉语字正腔圆、音调清晰, 英语则更加波纹起伏, 表现为流淌的、具有音乐性的语言。

2. **单音节与多音节。**汉语语素^②绝大部分由单音节构成, 也就是说大多数单音节词就可以准确对应一个含义。在这种语言传统下, 音节的持续时间会偏短, 音调轮廓也会在语句结束时迅速下降^③, 因此汉语出口往往顿挫有力、字字分明。而英语中占大多数的是多音节词, 像“u-ni-ver-si-ty”这样多达5个音节的单词也比比皆是, 因此英语表述会更为连贯, 声带震动的持续时间也往往更长。

3. **浊辅音的使用。**在现代汉语普通话中, 浊辅音仅有“m、n、l、r、ng”五个, 而英语则包含更多。这类音素发音时, 气流实际受阻和摩擦的位置靠下^④, 震动强度更大, 易产生更多的胸腔共鸣。如[b]、[d]、[g]三个音素, 在汉语中是不吐气的清辅音, 在英语中则是浊辅音, 试比较汉语中发“波”音的拼音字母 b, 与英语单词“book”中的[b], 需要声带发力的程度完全不同。

4. **元音的持续时间和口腔位置。**元音是指气流不受阻碍通过口腔发出的音。在这个过程中, 声道完全开放, 声门上方不会积聚气压阻碍气流传输, 所以元音通常携带了语音信号中的主体能量, 对最终的语态呈现有着极大影响。英语在表现元音上拥有巨大的优势, 这主要体现在两方面。第一, 英语中存在长短元音、单复元音的区别。声调不表义的语言特性导致元音持续时间成为了词义判断的重要依据, 这就要求英语在发音时必须充分、完整。为了不把 seat [si:t] (座位) 和 sit [sit] (坐) 混淆, 要特别注意元音时长的分别。而汉语作为声调语言, 并不存在依靠时间来表达语义的情况, 因此元音的总体持续时间要短于英语。第二,

^① 桂灿昆. 汉英两个语音系统的主要特点比较[J]. 现代外语, 1978(01): pp44-50.

^② 语素: 语言学术语, 是语音、语义结合的最小的语言单位, 主要功能是作为构成词语的材料。

^③ Levent M. Arslan, John H. L. Hansen. A Study of Temporal Features and Frequency Characteristics in American English Foreign Accent[J]. The Journal of the Acoustical Society of America, 1999, 102(1).

^④ 何善芬. 英汉语言对比研究[M]. 上海: 上海外语教育出版社, 2002.

相较于汉语，英语中的元音更多，并且更倾向于开口状态和后舌位（图 4.1.1）。这种元音发音时，口腔内部空间扩大，喉头位置降低，软腭抬高，参与震动的声带面积更大，相应的生理形态使得英语的元音听起来更加浑厚、饱满。

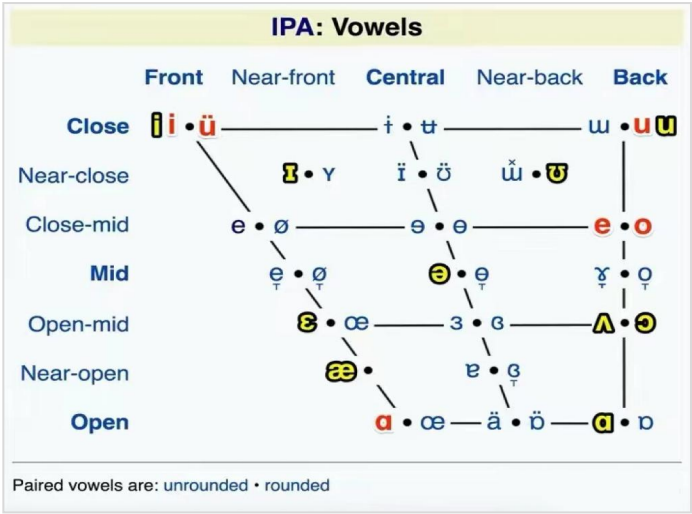


图 4.1.1 元音分布图^①

红色为汉语元音；黄色为英语元音。

横轴表示元音舌位前后；纵轴表示元音舌位高度，也可对应口腔开放程度；左右两侧表示发音时的圆唇程度。

二. 关于发音习惯

如果只是语言特征所致，为什么英语母语者在说其它语种时，往往也保持着磁性、低沉、浑厚的音色呢？笔者在查阅相关资料后发现，这并不是由人种因素导致的，实际上，不同人种平均下的口腔器官、声道、声带、各共鸣腔生理尺寸并无二致^②，这样的差异其实是口腔肌肉使用习惯长期累积的结果。也就是说，是语言习惯倒逼了特定发声器官的发育。

以汉语（普通话）、泰语为代表的汉藏语系发音时，力度集中在口腔前部，唇部肌肉和舌尖紧绷，发音辨识率高，但喉音偏重、音色单薄，与印欧语系的发音差异较大，而英语所属的正是印欧语系中的日耳曼语族。这类语言在发音时的紧张部位主要在咽腔、喉部和舌根，位置靠后，形成了与汉语不同形态的共振峰。如果以乐器比喻，汉语发音像是弦乐器，琴弦基础振动协同琴弓修饰，产生的声音具有金属感，而英语则更像管乐器^③，吐字饱满，发音圆润，共鸣厚重。

^① International Phonetic Association. IPA vowel chart[EB/OL]. 2005, <https://www.internationalphoneticassociation.org/content/ipa-vowels>.

^② Steve An Xue, Donald Fucci. Effects of Race and Sex on Acoustic Features of Voice Analysis[J]. Perceptual and motor skills. 2001, 91(3 Pt 1): 951-8.

^③ 张习如. 中文发音与英文发音的差异简述[EB/OL]. 2022.3.20, <https://zhuanlan.zhihu.com/p/32999560>.

三. 语言综合影响下的声音特征

在元音使用情况的影响下，英语母语者的低频泛音丰富，高频泛音稍弱。

在发音时习惯的低喉位、后舌位、高软腭形态的影响下，英语母语者的第一共振峰偏高，第二共振峰偏低。

在浊辅音、爆破音的使用下，英语母语者的胸腔共鸣更加丰富。

在发音时长、连贯度的影响下，英语母语者发声时的音腹更为饱满，持续时间更长。

四. 针对汉语的补充调整

在录制阶段，需要对演员的口腔状态和发声状态进行事先调整。放松唇部肌肉，打开口腔，提升软腭。在说台词时略微放慢语速，让每个音节的持续时间更长、发音更加饱满。整体的发声状态应弱化顿挫感，尽可能贴近英语发音所对应的流畅、连贯、粘性的特征。

在处理阶段，可以通过均衡器塑造 F1 略高、F2 略低的共振峰分布来模仿更开阔的口腔和更长的声道；可以通过提升低频谐波，削弱高频谐波的方式，模仿低喉位发声时，声带大面积参与振动形成的频率分布；可以通过短启动时间压缩音头，中等释放时间放出音腹的动态处理手段调整包络，弱化汉语字节之间的断裂感和颗粒感，使其更加具有连贯性；也可以直接了当地使用谐振滤波器添加共振，模仿英语发声时胸腔大量参与振动的感觉。

以上处理的目的是让共振成分不足的汉语语声，能在遵循本文宏观思路的基础上，依语种发声差异对汉语进行针对性的调整和补充，从而增强汉、英两种语声信号形态的近似性，这有助于汉语语声在本文的研究结论下获得更理想的效果。而关于如何能够跳出本文以英语为参考的创作思路，为中国电影创作出独属于汉语语言特色的语声形象，仍有待进一步研究。

第三节 “类人”语声处理思路应用在其它声音中的可能性

研究“类人”角色语声的处理问题，归根结底就是研究如何在保留原声音色特质的同时，表现出发声体所对应的特定外观形态和生理构造。这种创作思路的应用不仅局限于本文所论述的语声方面，也可以拓展至其它两类声音，即弱化处理用于常规人物对白、强化处理用于效果声音或怪兽声音。

人物语言是电影语言中表达情感最直接的途径，能够唤起观众最直接的心理

感受。对于大多数现实题材故事片而言，并没有创作特殊语声的需求，但这类思路同样可以为常规对白的音色处理提供参考。如果某现实角色有着较为标志性的外貌体征，或在故事中身份特殊，甚至是希望借由音色设计来表现角色间的情感对立、人物的内心纠葛，都可以根据具体需求选择性地使用上述“类人”角色音色的处理手段。巧妙的音色设计，可以让对白承载更多的人物特征和情绪氛围，具有更多创造或加强戏剧性的可能，给予观众一种超越表象的意蕴。

强化处理能够用于各类需要表现体积感、重量感的声音，并且选用手段和参数设置可以更加极端。以效果声和怪兽语声为例，这类声音对音色稳定性、自然流畅度的要求有别于普通对白，因此在处理时可以更加大胆。比如希望强调一扇石门的重量时，大尺度的变速、变调、加上激进的失真、调制、低频激励，有助于实现更加明显的效果。而怪兽语声则可以以本文“类人”语声的处理为基础，通过输入信号直接触发效果器内部样本，或借助声音融变手段将各类生物的吼叫声与语声素材融合，最终实现人化语声易懂度和兽化外型表现性的统一。

结 语

本文研究关注“类人”角色语声的创作，通过对人类发声机理、声学参数与体型的对应关系、中英两语种发声差异等原理的理性分析，以及在此基础上进行的一系列辅助实验，归纳出了多种手段的应用效果和适用倾向，希望能够丰富语声处理的相关研究，为实际的声音创作提供参考。

本文最终得出以下结论：


(1) 人体生理参数与声学参数的对应关系原理，对于角色语声形象的塑造有着重要的指导意义，在掌握原理的基础上进行针对性处理往往事半功倍。

(2) 语声创作的思路和规律，可以应用于由前期构思、语声录制、素材分析，直至后期处理的完整流程。以明确的语声形象为坐标，选择性地综合各种处理手段有助于获得更加理想的效果。

(3) 汉语、英语有着各异的语言特征和发音习惯，这两点共同导致了不同语种使用者的语声特征会在某些方面表现出差异。因此，仅借鉴英语适用的处理手段并不一定能使汉语达到同等的处理效果，应事先对汉语语声进行对应特征的补充。

受限于较大的学科跨度，本文的研究对象仅针对已完成的语声信号，对应的处理方式及选用软件也局限于影视、游戏和音乐制作领域。依托本文的理论基础和实验经验总结，能够在一定程度上帮助声音创作者们明确思路，进行分语种、分目的、分流程的调整，但仅以目前所选用的手段，终究无法实现对目标语声根源性的改变，这不仅限制了处理的精度和效率，更阻碍了创作者们对更加多样化音色的探索空间。如若未来与此课题相关的研究者可以将语音合成技术介入，从不同特质的语音信号如何生成的源头进行探究，或许会给电影特效声音的创作带来全新的思路 and 方向。


作者签署：








日 期：2023 年 5 月 14 日

附录（一）

《沙丘》混录师 Ron Bartlett 在邮件中的分享



Ron Bartlett 2022-10-22 12:39
发至 我

[详情](#)

Hello Ding Si Li,

Thank you so much for your kind words on our mix of "Dune". I'm flattered that you want to know all about how we created certain sounds and events.

"Dune" had so many layers of sounds and a LOT of design work put into it throughout the film. Some were very complicated and took a long time in the making and others were quite simple in the approach.

The Baron was part of the latter, meaning a much more simpler approach most of the time.

I used Fab filter EQ3 on all my tracks as well as a Fab Filter Compressor. Each scene varied in the reverb and delays used to match the appropriate rooms and environments. I used Altiverb, Cinematic Rooms, Stratus 3D and Slapper for a delay. Most of the sound of the Baron came from an aggressive EQ and compression with a lot of clip gain in protocols. I mix by the syllable and am very careful to dig out every nuance possible in the dialogue. Fortunately Sarsgard (the actor) has a very deep gravelly voice which gave me a lot to work with! I accentuated these tones in his voice with my EQ and brought up the presence of his timbre with the compressor.

In the scene where the silence barrier was used, I rolled off most of the top end with EQ and kept his dialogue just on the edge of perception to help that effect. I also used an appropriate reverb for the large room. Once inside the barrier, I used a series of short delays and panned them around the room to create a very claustrophobic environment. I wanted to replicate the sensation I had when I visited a temple that had a perfectly round room and you could hear a whisper in your ear very loudly because of the amazing acoustic properties.

When the Baron rises up and says "My Arakis, My DUNE!" is used different delays and reverbs in ATMOS to really accentuate his size and omnipotence.

I hope this gives you some insight on how I approached the Baron's dialogue. In the end it's all about getting a response from your audience and pushing the story forward with character.

I hope I had a small part in that to help Denis make a beautiful impactful film.

Let me know if you have any other questions I can answer.

Good luck on your paper!!

Ron Bartlett

参考文献

- [1] 桂灿昆. 汉英两个语音系统的主要特点比较[J]. 现代外语, 1978(01).
- [2] John A. Maurer iv. Research in Underwater Sound: With Focus on Musical Applications and Computer Synthesis[EB/OL]. 1998.3.15,
<https://ccrma.stanford.edu/~blackrse/h2o.html#hydrophone>.
- [3] Levent M. Arslan, John H. L. Hansen. A Study of Temporal Features and Frequency Characteristics in American English Foreign Accent[J]. The Journal of the Acoustical Society of America, 1999, 102(1).
- [4] Steve An Xue, Donald Fucci. Effects of Race and Sex on Acoustic Features of Voice Analysis[J]. Perceptual and motor skills, 2001, 91(3 Pt1).
- [5] 何善芬. 英汉语言对比研究[M]. 上海: 上海外语教育出版社, 2002.
- [6] International Phonetic Association. IPA vowel chart[EB/OL]. 2005,
<https://www.internationalphoneticassociation.org/content/ipa-vowels>.
- [7] 陈小平. 声音与人耳听觉[M]. 北京: 中国广播电视出版社, 2006.
- [8] T. Fitch. Encyclopedia of Language & Linguistics (2nd Edition)[M]. Amsterdam: Elsevier Science, Editor(s): Keith Brown, 2006.
- [9] M. Schroeder, Thomas Rossing, F. Dunn. Springer Handbook of Acoustics || The Human Voice in Speech and Singing[M]. NY: Springer New York, 2007.
- [10] 高蕾. 异形美学分析[D]. 西安美术学院, 2007.
- [11] Bernard J. Baars, Nicole M. Gage. Cognition, Brain, and Consciousness (2nd Edition)[M]. US: Academic Press, Editor(s): Bernard J. Baars, Nicole M. Gage, 2010.
- [12] Benjamin D. Charlton, William A. H. Ellis, Allan J. McKinnon. Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: Honesty in an exaggerated trait[J]. Journal of Experimental Biology, 2011, 214.
- [13] 王冬梅. 论声部划分的依据[J]. 才智, 2011, 10: 1.
- [14] Sheldon Schiffer. Footsteps, breath and recording devices: Abandoning a camera-centric construction of ‘point of audition’ [J]. The Soundtrack, 2012, 5(1).
- [15] Yi Xu, Albert Lee, Wing Li Wu. Human Vocal Attractiveness as Signaled by Body Size Projection[J]. PLoS ONE, 2013, 8(4).
- [16] 李晓辉. 汉语语声特点的分析及应用[D]. 北京邮电大学, 2014.
- [17] 孙兆. 汉语语声特征的研究[D]. 北京邮电大学, 2015.

- [18] 吴孟逸. 新好莱坞科幻电影外星人形象研究[D]. 西南大学, 2015.
- [19] 张悦. 计算机环境下人声合成技术的艺术化运用[D]. 南京艺术学院, 2016.
- [20] F.Alton Everest, Ken C. Pohlmann. 声学手册 (第 5 版) 声学设计与建筑声学实用指南[M]. 北京: 人民邮电出版社, 2016.
- [21] David Pietraszewski, Annie E. Wertz, Gregory A. Bryant. Three-month-old human infants use vocal cues of body size[J]. Proceedings of the Royal Society B: Biological Sciences, 2017, 284.
- [22] Stephan Reber, Judith Janisch, Kevin Louis Torregrosa. Formants provide honest acoustic cues to body size in American alligators[J]. Scientific Reports, 2017, 7.
- [23] Krotos LTD. Dehumaniser II User Manual[EB/OL]. 2017, <https://s3-us-west-2.amazonaws.com/dehumaniser/Manuals/Dehumaniser+2+Manual.pdf>.
- [24] iZotope Inc. VocalSynth 2 Help Documentation[EB/OL]. 2018, <https://downloads.izotope.com/docs/vocalsynth201/index.html>.
- [25] 汪月波. 米歇尔·希翁的电影声音理论研究[J]. 当代电影, 2018, No.266(05).
- [26] 席妍. 论人物语声在影视叙事中的艺术功能[J]. 文艺争鸣, 2018, No.290(09).
- [27] 杜士武. 咽喉机能原理与嗓音声学分析[M]. 吉林: 吉林大学出版社, 2020.
- [28] 徐泽南. 美国外星人电影的意识形态研究[D]. 湖南师范大学, 2020.
- [29] Waves Audio LTD. OVOx Vocal ReSynthesis User Guide[EB/OL]. 2020, <https://www.waves.com/1lib/pdf/plugins/ovox-vocal-resynthesis-v2.pdf>.
- [30] Jennifer Walden. Behind the Growing, Evolving Sound of ‘Stranger Things’ Season 3 - with Craig Henighan, Will Files and Mark Paterson[EB/OL]. 2020.7.10, <https://www.asoundeffect.com/stranger-things-season-3-sound/>.
- [31] 徐竟涵. 异星情境与创世想象: 雷德利·斯科特的异形美学[J]. 当代电影, 2021(08).
- [32] Marcus Gabriel. Dune Behind the Scenes: Creating Baron Harkonnen[EB/OL]. 2021.9.30, <https://dunenewsnet.com/2021/09/dune-behind-the-scenes-creating-baron-harkonnen/>.
- [33] Bryan Alexander. The making of oversized ‘Dune’ villain Harkonnen: no CGI, just a lot of prosthetics[N]. Chicago Sun-Times, 2021.10.29.
- [34] Katarzyna Pisanski, David Reby. Efficacy in deceptive vocal exaggeration of human body size[J]. Nature Communications, 2021, 12, 96812(1).
- [35] 刘艳. 科幻类电影非人类角色语音及非现实空间的声音设计特征研究[D]. 南京艺术学院, 2022.
- [36] How Does the Human Body Produce Voice and Speech?[EB/OL]. USA: National Institutes of Health, 2022.4.12, <https://www.nidcd.nih.gov/news/multimedia/how-does-human-body-produce-voice-and-speech>.

- [37] 王珏. 赋形、赋能与概念植入：中国魔幻电影声音符号建构[J]. 当代电影, 2022(10).
- [38] Yuyan Wei, Lin Gan, Xiangdong Huang. A Review of Research on the Neurocognition for Timbre Perception[J]. Frontiers in Psychology, 2022(13).
- [39] Brüel & Kjær. Hydrophone[EB/OL].
<https://www.bksv.com/zh/knowledge/blog/sound/gas-environment-hydrophone>.

致谢语

2021年8月，我在家中翻读完了那本厚厚的混音指南。当时有一章我怎么读都不理解，可面对庞杂的概念和知识，我不知道是否真的需要深究每处细节，也不知道如何继续才能摸到创作的脉络，于是斗胆给石宝峰老师发了一条信息请教。老师在打来的电话里给我讲解了近二十分钟，然后告诉我，忘掉他刚才教给我的。“没有必须要遵守的规则。不要怕，你大胆去做，大胆去尝试。”

感谢宝峰老师当年的这番话和他在601课堂上精彩的讲授，为我推开了那扇通往自由探索新世界的大门。在后来漫长的学习和论文写作过程中，我有过太多不解的、迷茫的时刻，幸好那些大大小小的困惑都能得到老师的耐心解答。在面对专业问题时，他会和我分享他多年积累的实践经验；在需要独立选择时，他会引导我讲出自己的想法，包容我遵循自己的意愿。我深知，知识是稍加时间就可以获取的东西，但独立思考的习惯和能力却远没有那么容易习得。感谢宝峰老师给予我的开放、平和、深刻，以及更重要的信任，因为这些，我才得以毫无顾虑地在这片奇妙而广袤的旷野中徜徉。

感谢 Ron Bartlett、辛海平、张新一为本文实验所做出的贡献，他们的帮助是让本文尽可能多地具有说服力的关键。同时，还要向本文引用文献的学者们和众多愿意在网络上分享他们宝贵经验的创作者们一并表示感谢，是他们的智慧成果为我的研究提供了坚实的理论基础和丰足的学习资源。

感谢录音系全体老师们对19级电影录音班的付出。感谢小江老师四年以来不遗余力的照拂，感谢龙老师为我提供的实践机会以及对我的帮助和肯定，感谢乐文老师、吴凌老师在研究生考试前夕对我的鼓励和支持。这些珍贵的情谊，我将一直铭记于心。

比起童年，青年可能更像是真正人生的开端，很开心青年时代启幕的这四年是和彼此珍惜的大家一起度过的。我们处在一切人对一切人的竞争中，只有爱是逃避这一切的港湾。

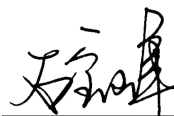
指导教师评语

该论文基于人体发声原理探究了如何在科幻类电影中进行类人型角色的语音制作。通过对具体片例、具体角色的分析和归纳总结，较好地基于理论原理探讨了特定类别声音的制作思路及制作方法，对于电影声音创作提出了个性鲜明的见解，具有一定的创作研究价值。

论文从多维角度做出了尽可能详实、严谨的论述。在该生掌握的相关知识范围内做到了灵活运用所学知识并结合实际创作加以验证，有效支撑了论述的科学性和严谨性。

建议丁思立同学在后续学习和工作过程中继续保持勤奋好学的态度和严谨扎实的作风，力争将研究成果早日应用于未来的实践创作当中。

教师签名：



日期： 2023 年 5 月 14 日