

# 北京青少年“女国音”现象：性别表演与社交媒体的影响

杨淇奥 田智中

2025 年 07 月 24 日

**摘要：** 本研究探讨了北京市青少年中“女国音”现象（即  $t\epsilon$  组声母前化现象）的使用情况及其背后的社会与语言动因。通过对北京市中学学生的语音样本、调查问卷与深度访谈数据进行分析，本文旨在回答两个核心问题：第一，现如今女国音的性别分布是否仍呈现明显的性别化特征；第二，传统的语言解释框架（如爱美心理说、保留古音说、亲密表现说）是否能解释当代女国音的变异模式。研究表明，尽管女国音仍主要在女性群体中使用，但男性的参与逐渐增加，打破了性别化的边界。此外，传统解释已无法充分解释女国音的当代变异，社交平台的语言模仿和“可爱”风格的传播对这一现象的扩散起到了重要作用。研究进一步探讨了女国音作为一种性别表演的风格化表征，揭示了性别规范在语言中的反复实践与变迁。此研究对性别语言学、语言变异及社会传播机制提供了新的视角，也为性别教育及语言规划提出了挑战。此外，本文还初步观察到了一种与“女国音”现象相反的，仅在男性说话人中出现的  $t\epsilon$  组声母后化现象，暂且命名为“男国音”。

**关键字：** 女国音； $t\epsilon$  组声母前化现象；性别表演；社交媒体；性别身份；媒介传播

# 目录

1 绪论 .....	3
1.1 研究缘起 .....	3
1.2 研究问题 .....	3
2 研究方法 .....	5
2.1 研究对象 .....	5
2.2 数据收集工具及过程 .....	5
2.2.1 语言样本收集 .....	5
2.2.2 调查问卷 .....	6
2.2.3 结构式访谈 .....	7
2.3 数据分析 .....	7
2.3.1 定量分析 .....	7
2.3.2 定性分析 .....	8
2.4 伦理考虑 .....	8
3 研究结果 .....	9
3.1 女国音使用情况 .....	9
3.2 tɕ 组声母前化变体使用者的自知性 .....	10
3.3 定量分析视角：社交平台分析 .....	10
3.4 定性分析：tɕ 组声母前化变体的使用与传播，一项性别表演的分析 .....	14
3.4.1 tɕ 组声母前化变体作为风格化的性别引述 .....	14
3.4.2 从网络到身体：性别规范的传播机制 .....	14
3.4.3 模仿、戏仿与界限测试：男生的前化变体实践 .....	14
3.4.4 声音的功能性选择：另一种非性别路径 .....	15
3.5 “男国音”：北京青年 tɕ 组声母的后化现象的初步观察 .....	15
4 总结讨论 .....	16
4.1 研究问题的回顾与回应 .....	16
4.2 主要发现 .....	17
4.3 理论与实践意义 .....	17
4.4 研究局限 .....	18
4.5 未来研究方向 .....	19
致谢 .....	21
参考文献 .....	22
附录 .....	24
附录 A：访谈文本（节选） .....	24
附录 B：数据分析代码 .....	30

## 一 绪论

### 1.1 研究缘起

早在上世纪 20 年代，黎锦熙先生就发现了北京地区女子中学的一部分学生会将声母 j, q, x (tɕ-, tɕʰ-, ɕ-) 读为 ts-, tsʰ-, s-, 取名为“女国音”。因为这一发音最先在原西城劈柴胡同的女子中学（今北京师范大学附属女中）出现，这种发音现象也被早期学者为“劈柴派读音”，不过现在已经没人这么称呼这种现象了。

在 1937 年，赵元任等人在高本汉《汉语音韵学研究》译本注中提及北京女学生普遍使用齿音的趋势(高本汉 等, 1940)。1950-70 年代，徐世荣又在有关北京话土音的文章中两次指出类似现象(徐世荣, 1956; 1979)。1980 年代，陈松岑 (1985) 在《社会语言学导论》中系统讨论了北京话女国音，指出其声母前移的特征，并将其归为 15-30 岁有文化女性的口音特征。随后曹耘于 1986-87 年发表了两篇专论（《北京话语音里的性别差异》《北京话 tɕ 组声母的前化现象》），报道了大量调研数据，对女国音的语音描述及社会成因展开讨论(曹耘, 1986; 1987)。同年，胡明扬发表了专题调查《北京话“女国音”调查》，这是迄今为止国内关于该现象最系统的实证研究之一(胡明扬, 1987)。进入 21 世纪，学界关注焦点逐渐扩展到语言与性别认同、媒体影响等社会语言学层面。2004 年，朱晓农从生物学视角提出“亲密高调论”，将女国音视为女性语言亲昵化倾向的体现(朱晓农, 2004)。

总的来说，全部研究都一致地认为女国音的主要特征是舌面辅音位置前移，即把原标准普通话中的 j, q, x 发音时舌面位置更靠前，读为 ts-, tsʰ-, s-。相应的，发音的时候口型会更小，送气更弱，导致音色听起来更加尖细。音韵上来看，女国音虽然把一部分 tɕ 组声母的字读为尖音，但女国音与历史上的尖团分音不同，并不区分尖音字与团音字，而是统一将尖音字和团音字都读为尖音(曹耘, 1987; 胡明扬, 1987)。

此外，在社会分布方面，女国音主要分布在受过教育、讲普通话或北京话的年轻女性中，而男性、儿童（青春期前）及中老年女性中出现率很低。胡明扬 (1987) 调查发现，使用女国音的女性并不自觉地意识到自己的发音不同，对应读词表时更倾向于使用标准普通话；一般在初一初二出现这一现象，婚后逐步减弱。

对于女国音的成因和传播方式学界众说纷纭。胡明扬通过问卷和实地调研总结了女国音的人群分布和社会成因，指出这一现象常见于非正式语域中。他认为，女性口型小和受声音尖细美感的影响使得 j、q、x 等声母前移（爱美心理说），曹耘也同意这个观点。但曹耘提出女国音可能受上海、苏州等地尖团历史区分影响的假说（保留古音说），胡明扬对此持怀疑态度(曹耘, 1987; 胡明扬, 1987)。朱晓农 (2004) 提出“亲密高调论”，认为女国音等高频声调现象源自婴幼儿语的亲昵化功能，是女性心理情感的象征表达（亲密表现说）。不同研究视角各有侧重，形成了关于女国音历史演变、审美动机、社会心理等多元争鸣的局面。

### 1.2 研究问题

尽管女国音在上世纪 80 年代到本世纪初曾受到学界短暂的关注，但自朱晓农 (2004) 之后，该话题在学术研究中几乎陷入沉寂，相关文献稀少，更遑论对女国音后续演变的系统性追踪了。这一沉寂现象的出现，既因为研究范式的变化，也与女国音自身的语言特性及其“非主流”的地位密切相关。

首先，从研究范式来看，2000 年代之后中国大陆社会语言学研究逐渐转向宏观语言政策、方言保护、语言接触与普通话推广等议题，而“小众变体”如女国音则被边缘化(Liu, 2022)。此外，虽然国际社会语言学中第三波变异研究强调性别与语言风格的交互构建(Podesva, 2011; 2007)，国内语言学研究的性别转向(gender turn)尚处早期萌芽阶段，缺乏对性别与语言交互构建的深层理论关怀，使得女国音难以获得持续的理论关注。其次，女国音本身作为一种非标准的发音变体，常常被视为临时的语言时尚或边缘社群的语音偏差，很难引发主流语言学界对其语言系统性、结构层级的兴趣(Milroy, 2002)。相较于具有区域方言系统或民族语言背景的变体，女国音缺乏正当的语言身份，因此在研究资源有限的情况下被优先级低估(Coupland, 2001)。

综上，2004 年后女国音陷入沉寂并非因为其在现实中消失，反而是在缺乏理论工具与方法创新的前提下，这一不断发展的语音现象被研究者“视而不见”。也正因如此，本文才希望在性别语言学、媒介语言学的交叉点上，重新激活这一议题的研究生命。本研究试图在前人的基础上，重新审视这一语言现象，聚焦于以下两个问题：

第一，现如今女国音的性别分布是否仍延续胡明扬 (1987) 所描述的格局？

胡明扬 (1987) 曾指出，女国音主要出现在都市中受过较好教育的女性话语中，其本质是一种“趋美心理”的语言体现；而曹耘 (1987) 则强调其与京剧传统、南方方言影响以及“老国音”的残余有关。然而，三十余年过去，北京方言的社会构成、传播媒介与性别意识形态皆已发生深刻变化。

本研究关注的第一项议题，正是女国音在当代青少年语音实践中的性别分布是否仍呈现明显的性别化倾向。换言之，是否如胡、曹所述，依然是女性话语中的主导特征，抑或这一现象已向更多性别群体扩展，甚至在部分男性话语中被模仿、采纳，形成某种“反转”的性别语言实践？此外，性别与社会身份之间的交互作用是否影响发音选择？这些问题将通过对几十位初高中学生的语音材料，访谈资料及调查问卷加以分析。

第二，女国音的语言动因是否仍符合胡明扬 (1987)、曹耘 (1987) 的解释框架，抑或已呈现新的动力机制？

本项研究问题关注的是语言行为背后的社会动因。曹耘 (1987) 认为女国音的形成受南方方言迁入、京剧艺术影响和国语运动时期“老国音”残留所驱动，而胡明扬 (1987) 则强调“美音”心理对个体语言行为的塑形力量。这些解释虽然在特定历史语境中成立，但是否仍能充分解释当代女国音的再生产，显然值得质疑。

本研究尝试通过深度访谈，探讨受访者对女国音这一现象的主观理解与态度。社交平台的模仿机制也可能助长了这一语音特征的扩散。因此，除了回顾“老国音—京剧—南方话—爱美心理”四点模型外，本文还将考虑新的动力机制是否正在逐步取代传统解释框架。

以上两个研究问题将作为本研究的核心出发点，贯穿随后的语料分析与理论讨论部分。随后在第二章中将进一步详细阐述研究所采用的方法，包括语音采集方式、访谈对象选择标准以及分析路径设定。

## 二 研究方法

本章将详细介绍本研究的设计、参与者招募、数据收集过程等。为确保研究的严谨性，本研究采用了定量与定性相结合的研究设计，通过问卷调查与深度访谈相结合的方式，系统地探讨北京市中学生中的女国音现象。

### 2.1 研究对象

本研究的参与者来自北京市不同学校的中学生。参与者总共为 44 人，其中 17 名学生来自北京市朝阳区外国语学校初中部（7-8 年级），27 名学生来自北京市第八十中学望京校区（9-12 年级）。所有学生均为随机选择，保证样本广泛并具有代表。研究过程中，我们在校门口、校园内和食堂处随机询问了这些学生，确保他们同意参与调查后才开始了后续调查。

### 2.2 数据收集工具及过程

本研究设计了三种数据收集工具：语言样本收集、调查问卷和结构化访谈。

#### 2.2.1 语言样本收集

我们首先要求参与者朗读一段故事短文。短文是基于《北风与太阳》故事加以修改而成的，特别加入了多个以 j, q, x 为声母的字，方便观察参与者在连续语境下的发音特征。我们使用的短文如下，其中以 j, q, x 为声母的字标有下划线：

有一次，北风和太阳正在争论谁的本领强。他们看见一个行人穿着一件夹克，在街道上匆匆前行。他们就说，谁可以让行人脱掉那件斗篷，就算谁比较**有本事**。北风急着展示自己的力气，呼呼吹起了狂风。怎料，行人紧紧抓住衣角，把衣服裹得更紧。最后，北风没办法，只好放弃。接着，太阳不急不躁，轻轻地照着大街，气温渐渐升高。行人开始觉得热，不由得解开了夹克。于是，北风只好认输了。

此外，我们还要求参与者朗读一个词表。因为胡明扬 (1987) 使用的词表制作精良，包含历史上本来的尖音和团音，也考虑到了不同介音、不同开口程度的字，我们直接采用了这篇文章中的词表，如表 1 所示。

表 1 词表

1	2	3	4
鸡	成绩	信心	交际
急	积极	新鲜	消失
几	整齐	王府井	舅舅
寄	机器	安静	教育局
漆	希望	九斤	秋千
棋	看戏	清明	曲线
起	街道	高兴	休息
气	先生	行不行	需要
稀	茄子	千万	穷人
席	接近	坚决	英雄
洗	谢谢	前门	绝对
细	有一些	陷害	学习
		抢先	平均
		详细	安全
		将军	宣传

本研究和既往研究不同，即使用文章，又使用词表。这是因为胡明扬 (1987) 提到，一些北京人在读文章时相较读词表时更紧张，所以读文章和读词表时可能一个会出现女国音，另一个不会。但他并没有证明这一点，只是这样推测。所以我们借助上述文章和词表同时使用，正好可以验证是否如此。

## 2.2.2 调查问卷

调查问卷主要包括以下几部分内容：

1. **方言背景：**参与者需填写个人的方言背景信息，内容包括：是否在北京出生、长大；父母的籍贯；家庭中使用方言的情况；是否能使用其他方言等。方言背景作为影响个体语音认知与产生的核心社会变量，已被多项研究证实在语言变体的形成与再现中扮演关键角色(Labov, 1972; 张振兴, 2013)。尤其在有无“尖团音对立”的母语环境下，个体对 tɕ 组声母的知觉灵敏度与产出模式可能存在显著差异(曹耘, 1987)。因此，本研究特别考察受访者是否具备来自江南、华南或晋语等分尖团方言区的语言接触史(中国社会科学院语言研究所 等, 2012)。
2. **普通话与女国音的评价：**为了了解学生对不同语言风格的认知与态度，问卷中包含了关于普通话与女国音的评价问题。参与者需要根据 5 点量表（1 表示“非常不自然”到 5 表示“非常自然”）选择自己在日常生活中听到普通话和女国音的频率、习惯程度。此外，参与者还需评估男生和女生使用普通话与女国音的自然程度，并判断两种发音方式是否相同。个体对语言风格的认知态度直接关联其语言选择与社会认同建构(Eckert & McConnell-Ginet, 2013)。Butler (1990) 指出，语言作为性别表演的手段，其“自然性”是社会规范灌注的产物。通过评价普通

话与女国音的自然程度与适用性，我们得以观察被访者是否将前化发音与特定性别标签或身份期望绑定。

3. **社交软件使用情况：**问卷中还询问了学生常使用的社交软件、观看内容类型以及认为自己是否受到短视频博主语言风格的影响。社交媒体使用频率与内容偏好越来越被视为影响语言变体传播的新兴社会变量(Androutsopoulos, 2014)。多项研究表明，网络语言模仿是语言风格再生产的重要机制，尤其在青少年群体中更为显著(Lee, 2016)。在多模态平台（如小红书、抖音）中，特定语音风格（如尖细、高调、软萌）常被作为“流量语言”被传播和模仿，这为我们理解女国音的持续传播提供了新的解释框架。

问卷的设计旨在获取参与者对语言现象的态度、社交媒体语言使用习惯以及方言背景，从而为后续的数据分析提供有力支持。

### 2.2.3 结构式访谈

在问卷调查后，本研究随机选择了9名北京市第八十中学的学生和1名北京市第八十中学的年轻教师（5位男性、5位女性），这些学生分别包括有女国音特征与没有女国音特征的个体。访谈的目的是深入了解学生对于女国音的看法与态度。

访谈中未直接提及“女国音”这一术语，而是通过描述与示范这一发音现象的方式，让受访者回答他们认为这种发音是从何而起的，并询问他们对这种发音的态度。访谈问题主要包括：有没有注意到身边有人有女国音现象？对于女国音的态度是怎么样的？认为可能的传播方式是什么？

访谈数据的目的是在问卷数据基础上，通过受访者的主观反馈，帮助分析和解释其背后的社会文化动因，包括审美倾向与语言意识形态、性别规范与社会期待、网络模仿与平台风格、社交策略与情感动机和传播环境与媒介语境。

## 2.3 数据分析

本研究的数据分析方法包括定量分析和定性分析相结合。

### 2.3.1 定量分析

定量分析分为对参与者语音的定量分析和调查问卷结果的定量分析。

对于参与者使用女国音的程度，我们和胡明扬 (1987)一致，将其划分为三类。一类是完全不使用女国音，即使用  $t\epsilon^-$ ,  $t\epsilon^h-$ ,  $\epsilon^-$ ，我们称其为**标准发音**。一类是完全前化成龈腭音，即使用  $ts-$ ,  $ts^h-$ ,  $s-$ ，我们称其为**完全前化**。另一类介于两者之间，即使用轻微的前化，发音为  $t\epsilon_{\uparrow}^-$ ,  $t\epsilon_{\uparrow}^h-$ ,  $\epsilon_{\uparrow}^-$ ，我们称其为**半前化**。后两类属于女国音的范畴。

判定一份录音是否具有女国音特征时，我们采用感知性语音判断（perceptual phonetic judgment）的方法，结合发音者的齿龈擦音表现与整体音色特征进行判断。具体来说，当  $t\epsilon$  类声母在发音中呈现出持续性强、气流集中、音质嘶亮的特征时，即可判断其存在前化倾向。在操作上，我们



参考既有文献对齿龈擦音的感知特征描述(Ladefoged & Johnson, 2015)，将高频嘶嘶声(s-like hiss)作为前化的主要判断依据。在此基础上，将样本划分为标准发音、半前化与完全前化三组。本研究判断过程中未使用 Praat 等语音分析软件进行声谱测量，而是依托研究者对普通话音系的熟悉程度与语音变体的感知经验进行归类。这种方法虽带有一定主观性，但在具备语言学训练背景的听者之间仍具有较高一致性，并已被广泛应用于社会语言学与音变研究中(Labov, 1982)。

此外，通过调查问卷中的五点量表数据，采用描述性统计分析学生对普通话与女国音的认知态度及其社交软件使用情况进行量化分析。

### 2.3.2 定性分析

定性分析部分，本研究主要通过对结构式访谈数据的内容分析法(content analysis)来提取学生对女国音现象的看法和态度。研究者对访谈记录进行编码和主题归纳，分析受访者对女国音的来源、意义和性别差异的认知。之所以选择内容分析法，是因为该方法适用于处理结构化程度较高的访谈数据，能够在保留语境信息的同时，系统归纳受访者言语中反复出现的关键词、话语单元与主题模式(Krippendorff, 2019)。

与话语分析或叙事分析等方法相比，内容分析更关注语义内容本身而非语言形式或讲述顺序，特别适用于本研究所关注的“语言态度”“性别认知”“传播路径”等认知类变量的归纳分析(Bloor & Wood, 2006)。在操作上，我们采用开放式编码结合聚类归类，对受访者言语中涉及“模仿来源”“使用动机”“性别联想”等语段进行标记、归纳，并最终提炼为若干高频主题，以支持后续理论分析。

## 2.4 伦理考虑

本研究严格遵循伦理审查与隐私保护原则，确保所有参与者的权益不受侵犯。在数据收集过程中，所有参与者均为中学生，研究团队通过教师中介方式向学生明确说明研究目的、内容与数据使用范围，确保其在充分理解的基础上自愿参与，并享有随时退出的权利。

鉴于本研究为非干预性的语言行为观察，且所有数据均为匿名采集，不涉及姓名、年龄、家庭籍贯等可识别身份信息，调查问卷与语言样本中亦未设置任何敏感内容，因此研究过程中未直接联系家长征求书面同意，但已确保参与过程在教师监护与告知下进行，符合法律与伦理的最小风险原则。

本研究采取了以下具体保护措施：

- 问卷调查不要求填写姓名，所有身份信息不在论文中披露；
- 语言样本（如读词表、读文章录音）仅用于研究内部分析，不对外公开；
- 访谈录音不披露，转录文本已进行匿名处理，仅在论文中引用和附录内容经过脱敏的内容；
- 所有访谈问题均以中性方式表述，避免引导性提问与任何可能引发不适的术语。

通过上述伦理设计，研究力求在保障数据有效性的同时，最大限度地保护参与者的隐私权、知情权与表达自由。



三 研究结果

3.1 女国音使用情况

在本研究调查的 44 人中，女学生占 21 人，年级分布从七年级到十二年级；男学生占 23 人，年级分布从七年级到十一年级。读文章时，7 名女学生和 1 名男学生出现了半前化，分别占各性别总人数的 33.3% 和 4.3%，不过没有人在读文章时使用了完全前化。读词表时，11 名女生和 6 名男生说话时使用了女国音，分别占各性别总人数的 57.1% 和 26.1%；其中有 3 名女生和 1 名男生出现了完全前化的现象，剩下 9 名女生和 5 名男生使用的是半前化。按照年级分类的详细使用情况参见表 2。

表 2 女国音按年级、性别的使用情况<sup>1</sup>

年级	总人数		文章前化人数		词表前化人数	
	男	女	男	女	男	女
七	5	4	0	1	0	3(1)
八	6	4	1	3	2(1)	3(1)
九	2	4	0	1	0	2
十	6	2	0	1	3	1
十一	4	5	0	1	1	2(1)
十二	0	2	0	0	0	0

胡明扬 (1987)调查发现，在调查的 65 名男性中，无一人使用女国音，说明在该取样范围内，女国音属于女性专属的语言特性。而我们的调查结果显示，仅仅是调查的 23 名男性中，就有 6 名男性使用不同程度的前化，占总体的 26.1%。这与曹耘 (1987)调查中提出的 45 名男性中 13 人前化的比例 (28.9%) 十分接近。

虽然调查数据仍然显示，使用女国音的还是女性居多，但女国音现象很大程度上已脱离了其“女”的限制，在男性中也有一定传播。这也就是为什么我们认为女国音已经不再适合叫“女”国音了。我们在这里呼吁大家使用更为中性的称呼，如“tɕ 组声母前化现象”（简称“前化现象”）等。本文接下来也将使用“tɕ 组声母前化现象”或“前化现象”来指代女国音现象。

值得注意的是，在读文章时出现半前化的学生均在读词表时出现了不同程度的前化，而读文章时没有前化的学生在读词表时也有可能出现半前化。这也印证了胡明扬 (1987)提出的假设，认为读词表更能反映 tɕ 组声母前化现象的实际使用情况。

我们发现，在研究中发现使用前化现象（半前化+完全前化）的这 18 人里，虽然有 3 人在家中使用方言，6 人会讲方言，占前化人数的三分之一，但他们所使用的方言均来自华北、东北、西北，使用的方言很大程度上不分尖团音。即使有人父母来自分尖团的方言区（如一名九年级学生，父亲来自安徽，而皖北部分方言分尖团音），但这样的学生在家里并不使用方言，也不会讲方言，基本可以排除来自母语的影响。

<sup>1</sup>表中括号中的绿色字表示使用完全前化的人数。因为使用半前化人数显著多于完全前化，就不再单独标注了。

### 3.2 tɕ 组声母前化变体使用者的自知性

此前多篇文章都曾提到，前化变体使用者一般不自知(曹耘, 1987; 胡明扬, 1987)。也就是说，使用前化变体的人一般不认为自己说话和别人有什么不同，而是默认的使用了这样的变体。我们的调查也印证了这一点。在调查问卷中，我们特意包含了这样一个问题：你感觉你更常用哪种发音方式？是发音方式 1（不前化，由调查人员朗读部分词表，现场演示）还是发音方式 2（前化，同样由调查人员演示），抑或是认为两种发音一样，难以区分？调查结果详见表 3。

表 3 对于自己前化变体使用的认知

前化使用情况	认为一样	认为更常用标准发音	认为更常用前化发音
前化发音使用者(共 17 人)	9(53%)	6(35%)	2(12%)
标准发音使用者(共 27 人) <sup>2</sup>	9(33%)	12(44%)	5(19%)

调查结果显示，超过一半（53%）的前化变体使用者不能区分不前化发音和前化发音，自然也很难意识到自己正在使用前化变体。更有超过三成（35%）的前化变体使用者即使在读词表和文章时或多或少的使用了不同程度的前化发音，但认为自己更常用不前化发音。这项统计证明，有 88% 的前化发音使用者都不能准确的判断出自己的发音属于标准发音还是前化发音，更加印证了之前文章提到的前化变体使用者不自知的观点。此外，我们还发现有三成（33%）的标准发音使用者分不清两种发音，19% 的标准发音使用者错误地认为自己更常用前化发音。

这样的不自知性有可能是因为日常语言风格的自动化习得导致的。发音习惯是在长期社交互动中潜移默化习得的，而这样的前化发音可能是受到身边同伴或网络博主无意识影响，并非刻意模仿或选择。因此，语言风格的使用未必等于意识到自己选择了某种发音，一部分人可能真正难以分辨出自己的发音到底属于哪一类。

而不自知性另一种原因是社会评价驱动下的自我异化。简而言之，在互联网环境下，有些人即便使用了前化，也可能因为这样的发音被贴上矫揉造作的标签，而下意识否认自己使用了前化发音。这是一种避免被社会贴标签的语言防御机制。这类说话者并非不知道怎么说，而是不愿承认自己“像她们那样说”。

### 3.3 定量分析视角：社交平台分析

正如 小节 2.2.2 所述，我们使用调查问卷收集了 44 名学生常用的社交软件、在社交软件上常看的内容以及受试者受博主说话风格影响大小的自我评价，其中 17 人使用前化变体，27 人不使用前化变体。我们使用 Python 分别从以下几个方面分析了这些数据：社交媒体使用频率（各平台使用人数占比双比例检验），内容偏好频率（各类别观看人数占比双比例检验），说话风格受影响程度（5 级量表均值比较），组别与社交媒体使用的卡方独立性检验和其他补充检验，如影响程度 Mann-Whitney U 检验、社交软件使用数量 t 检验和内容类型观看数量 t 检验。

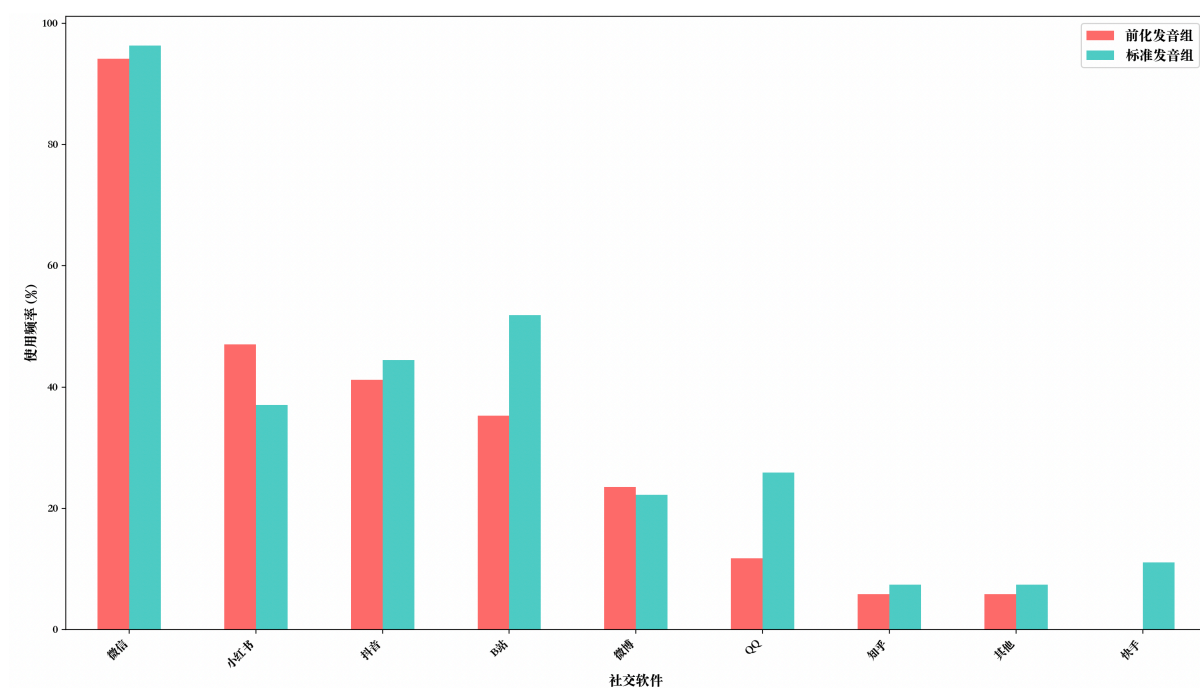
<sup>2</sup>有一名非女国音使用者在问卷上没有作答这一问，因此表格第二行总数加起来并不是 27。

我们首先分析了各社交平台使用人数占比。

表 4 各社交平台使用率双比例检验结果

社交软件	前化变体组使用率(%)	标准发音组使用率(%)	z 统计量	p 值	显著性
快手	0.0	11.1	-1.4238	0.1545	否
QQ	11.8	25.9	-1.1339	0.2568	否
B 站	35.3	51.9	-1.0740	0.2828	否
小红书	47.1	37.0	0.6583	0.5103	否
微信	94.1	96.3	-0.3378	0.7355	否
抖音	41.2	44.4	-0.2131	0.8313	否
其他	5.9	7.4	-0.1954	0.8451	否
知乎	5.9	7.4	-0.1954	0.8451	否
微博	23.5	22.2	0.1007	0.9198	否

图 1 社交软件使用频率对比



可以看到，微信在前化变体使用者和标准发音使用者两组人中渗透率都最高，分别达到94.1%和96.3%，这与该平台作为基础通讯工具的市场地位相符。次要社交平台则差异较大。垂直对比两组使用的社交软件，前化变体组更倾向使用小红书（47.1% v.s. 37.0%）和微博（23.5% v.s. 22.2%），而标准发音组在抖音（41.2% v.s. 44.4%）、B站（51.9% v.s. 35.3%）和QQ（25.9% v.s. 11.8%）的使用率更高。尽管这些差异未达统计显著性（所有  $p > 0.15$ ），但其绝对差高达12-16个百分点。这样的差异幅度值得我们关注。

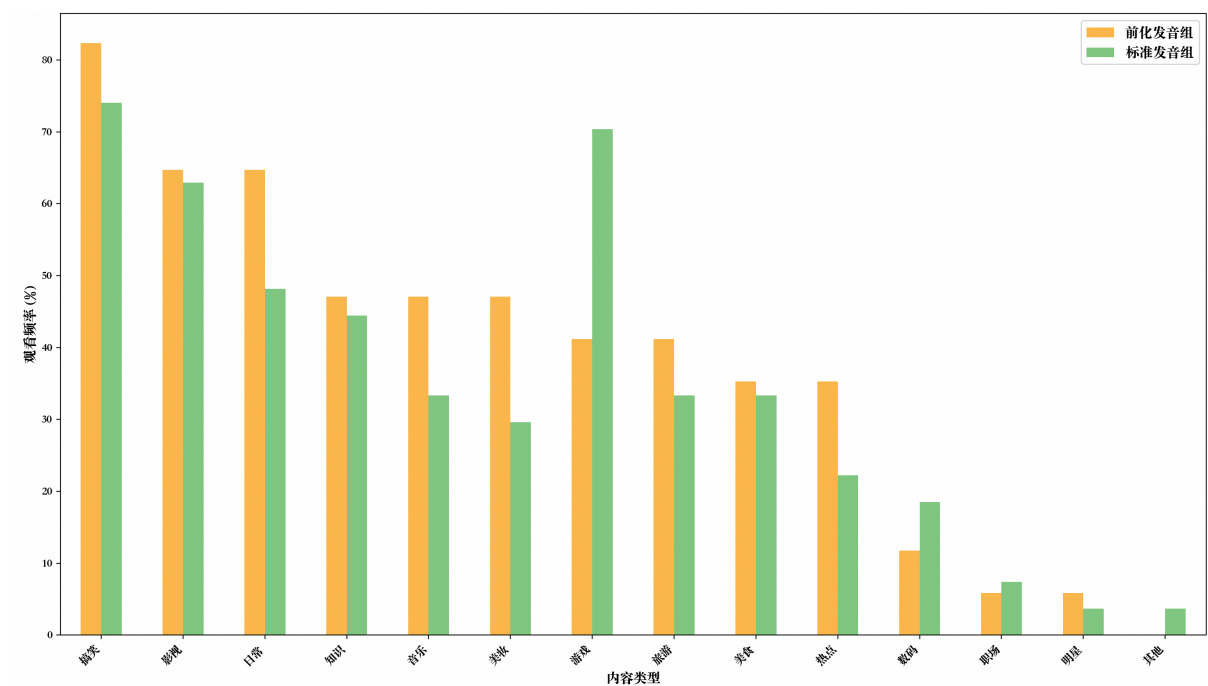
对比平台使用数量，前化变体组平均使用2.65个平台，而标准发音组则接触更广，平均使用3.04个平台。但这样的差异仍然未达到显著性（ $t$  统计量=-1.0827， $p$  值=0.2851）。

在调查问卷中，我们提供了如下几个选项，基本上涵盖了社交平台上各种类型的内容，受试者可以多选：搞笑/段子、日常 Vlog/情感倾诉、美妆/穿搭/变装、知识科普/冷知识、美食/探店/吃播、游戏实况/搞笑游戏、音乐/舞蹈/翻唱、影视混剪/解说、数码科技/评测、旅游风景/旅行攻略、热点新闻/社会评论、职场成长/教育分享、其他（可补充）。完成调查后，我们分析了受试者在不同社交平台上观看的内容类型。

表 5 不同内容类型观看率双比例检验结果

内容类型	前化变体组观看率(%)	标准发音组观看率(%)	z 统计量	p 值	显著性
游戏	41.2	70.4	-1.9178	0.0551	边缘显著
美妆	47.1	29.6	1.1702	0.2419	否
日常	64.7	48.1	1.0740	0.2828	否
热点	35.3	22.2	0.9480	0.3431	否
音乐	47.1	33.3	0.9104	0.3626	否
其他	0.0	3.7	-0.8027	0.4222	否
搞笑	82.4	74.1	0.6381	0.5234	否
数码	11.8	18.5	-0.5964	0.5509	否
旅游	41.2	33.3	0.5266	0.5985	否
明星	5.9	3.7	0.3378	0.7355	否
职场	5.9	7.4	-0.1954	0.8451	否
知识	47.1	44.4	0.1696	0.8653	否
美食	35.3	33.3	0.1336	0.8937	否
影视	64.7	63.0	0.1170	0.9068	否

图 2 不同内容类型观看频率对比



结果显示,大部分内容类型在前化变体使用者和标准发音使用者间并没有显著差异。虽然游戏实况/搞笑游戏这一类内容在两类人群中的差异并不严格地具有统计显著性,但  $p$  值 (0.0551) 十分接近显著性判定的标准 ( $p < 0.05$ ), 我们判定它具有边缘显著性。因此, 游戏实况/搞笑游戏的内容被更多的标准发音使用者浏览。

这一结果虽为局部差异, 却可能揭示出媒介内容消费与语言风格之间潜在的关联机制。虽然大多数内容类型在前化变体使用者与非使用者之间差异并不显著, 但值得注意的是, 美妆/穿搭/变装 ( $p = 0.2419$ )、日常 Vlog/情感倾诉 ( $p = 0.2828$ )、热点新闻/社会评论 ( $p = 0.3431$ ) 以及音乐/舞蹈/翻唱 ( $p = 0.3626$ ) 等内容在前化变体使用者中呈现出更高的偏好趋势。尽管未达到显著性标准, 这些倾向性差异在方向上一致, 且与前化变体普遍关联的社会意涵——如精致化、自我表达、情感可见性等——相呼应。

与此相对, 在“游戏实况/搞笑游戏”这一类中, 标准发音使用者的观看比例显著更高。此类视频多由男性主播主导, 语言风格通常更为粗犷、直接, 表达方式亦更贴近“阳刚”或“对抗性”的语气氛围, 这与前化变体的语音特征构成明显对照。因此, 非使用者偏好此类内容, 可能意味着他们在语言建模过程中更易接触并模仿这类 UP 主的非前化发音风格, 从而在现实交流中回避前化变体所体现的音系特征。

这一趋势亦可在社交平台的使用偏好中得到呼应。前文提到的平台使用频率的差异值得我们关注, 因为不同平台主导的内容生态与语言风格关联紧密。例如, 小红书与微博强调精致生活、自我叙述与情绪表达, 平台博主多采用带有“软化”“可爱”“亲昵”语感的表达方式, 这可能强化用户对前化变体的接触与模仿。而 B 站与抖音等平台, 尤其在游戏、鬼畜、搞笑等子频道中, UP 主语言风格更具攻击性、节奏快、缺乏前化, 使用者长期沉浸其中或更倾向于认同该类语言风格。

此外, 这种内容偏好差异也可能反映出媒介接触与语言风格、社群认同之间的互动关系。游戏类视频的受众往往具有特定的次文化归属感, 其主导的表达模式与言语习惯可能强化对“女性化”语言的回避倾向, 从而在一定程度上影响观众的语言表现。基于此, 可以将媒介内容类型的选择视为一种间接但持续的语言社会化过程。

进一步地, 从 Judith Butler 的性别表演理论 (gender performance) 的角度看, 个体在不同语境下的媒介消费与语言选择, 实则共同参与了性别角色的动态建构 (Butler, 1999)。在这一过程中, 语言风格 (如本文所讨论的前化变体) 不仅是语音层面的变异, 更是青年群体通过媒介行为与表达行为交互塑造的社会风格之一。因此, 尽管整体差异有限, 但游戏实况/搞笑游戏偏好的显著性差异, 仍为理解前化变体使用者与标准发音使用者之间潜在的语言认同差异提供了重要线索。

我们最后进行了影响程度 Mann-Whitney U 检验、社交软件使用数量  $t$  检验和内容类型观看数量  $t$  检验。这三种检验结果在前化变体使用者与非使用者间没有显著差异, 说明前化变体的使用与受访者的社交媒体内容偏好数量、使用平台的数量以及其自我感知受到博主语言风格影响的程度之间, 并不存在统计学上显著关联。



### 3.4 定性分析：tɕ 组声母前化变体的使用与传播，一项性别表演的分析

在本研究的定性分析中，笔者尝试借助 Judith Butler (1990, 1993) 提出的性别表演理论 (gender performativity)，对受访者关于 tɕ 组声母前化变体现象的描述进行分析。Butler 认为，性别并非一种预设的本质，而是个体在特定社会规范与规训框架下，通过反复行为（包括语言、姿态、服饰等）所表演出来的社会身份 (Butler, 1999)。在此理论视角下，前化变体不再仅仅是一个语音变化，而是一种“女性气质”的风格化表演，一种经由语言实践不断重复、巩固乃至变异的性别身份构建方式。

#### 3.4.1 tɕ 组声母前化变体作为风格化的性别引述

从多个受访者（如受访者 T、Z、X、O、L）的叙述中可以看出，前化变体并非自然习得，而常常被归因于模仿特定的声音模范，如“网红”“台湾口音”“译制腔”或“郭老师”式短视频话语风格。例如 T 提到，“可能是有些网红他们说话特别怪……别人就去 copy”；Z 和 X 明确指出台湾口音和可爱感之间的联系：“[我觉得前化变体这种说话方式]可爱。因为我追星追台湾人，我觉得他们说话很好听啊”。这些模仿与再现恰恰印证了 Butler 所谓的引述性 (citationality)：性别表现并非源自主体的内在真实性，而是对文化中已存在的性别样式的引用 (Butler, 1999)。

更重要的是，这种模仿在语境中具有重复性、可识别性，构成了 Butler 提到的风格化重复 (stylized repetition) (Butler, 1999)。这些重复行为将一种本属局部社群的语言变体，扩散为青年群体中被识别为女性化表达的语言风格，在日常互动中被认定为“可爱”乃至“装腔作势”的说话方式。

#### 3.4.2 从网络到身体：性别规范的传播机制

受访者 O、L 等受访者的描述揭示了前化变体的传播与网络语境密切相关。他们指出“网络影响挺大的”。受访者 O 明确表示模仿行为“是为了拉近关系”，在互动中下意识地完成。也就是说，前化变体不只是网络模仿的语言特征，它还是一种通过身体、声音、言语共同施行的性别位置入场仪式。

这种表演并非全然出于主观自由选择，而是在社会框架（如“女生要温柔”、“说话要软”、“可爱即正义”）中被召唤出来的行为模式，正如 Butler 所述：我们成为性别，是因为我们不断表演它，并由社会他者的反馈确认其有效性 (Butler, 1999)。这种反馈往往带有明确的性别价值取向，正如受访者 X 所说：“比较可爱，就会改成这样[说话]”，这实际上是对“合格女性”声音的一种审美评判。

#### 3.4.3 模仿、戏仿与界限测试：男生的前化变体实践

在前面的调查中，我们发现在调查样本中有 28.6% 的男性会使用 tɕ 组声母前化变体，而受访者 L 正是其中之一。男生使用前化变体这一现象恰恰印证了 Butler 关于性别并无原型的观点：当男生使用女性化的语言风格时，性别身份的表演性便显现出来。这种模仿既可能是戏仿性的（比如搞笑、破圈、嘲弄），也可能具有酷儿化的潜能，挑战了声音于性别绑定的自然化假设 (Butler, 1999)。

正如 Butler 所言，性别规范之所以具有稳定性与强制力，并非源于其本质上的合理性，而是因为它通过社会实践中的不断重复，使其“看起来”仿佛是自然的、生理的，从而掩盖了其文化建构的本质



(Butler, 1999)。而颠覆这种性别规范的关键在于重复中出现的错位：当前化变体由男性发出，其性别标签的不确定性本身就构成了对于规范性秩序的扰动。这种扰动未必总是反抗性的，也可能是自嘲、幽默，但正是这些“模仿的模仿”让我们意识到所谓“女性声音”，不过是被社会不断训练、确认并习以为常的一种文化表演。

#### 3.4.4 声音的功能性选择：另一种非性别路径

然而，值得注意的是，并非所有受访者都从性别视角理解  $t\epsilon$  组声母前化变体。例如，受访者 J（30 岁，女，教师，完全前化）提出了一种功能性的解释路径。她强调音质的清晰度和抗噪性。她认为 /s/ 更容易被听清楚，而 / $\epsilon$ / 会爆音，而她在录制课程时会主动避免使用 / $\epsilon$ /。这一解释可借助刘雅琦等 (2024) 的声学研究。他们从感知正确率、声学特征与抗噪性能等角度，对普通话里的 /s/ 与 / $\epsilon$ / 的“感知强度”作过直接比较。在安静条件下，/s/ 的总体识别正确率为 71.5 % - 83.3 %，而 / $\epsilon$ / 仅为 66.7 % - 68.3 %，/s/ 显著高于 / $\epsilon$ /。这一差距在噪声条件（-4 dB 与 -8 dB SNR）下进一步加大，说明 /s/ 的“感知强度”更强、抗噪性更好。另外，在 -i、-a、-u 三种元音前，/s/ 的感知正确率均高于 / $\epsilon$ /；且 -i 环境下两者的差距最大。造成这些差距的声学原因是：/s/ 的谱重心更高、能量更集中。/ $\epsilon$ / 的频谱相对平坦，容易被噪声掩蔽。受访者 J 的解释体现了一种以传播效果为主导的感知优化策略，而非性别表演。

在此框架下， $t\epsilon$  组声母前化变体并非性别施行，而是一种针对媒介环境的实用性调音。这种观点挑战了过度符号化声音风格的趋势，提醒我们语言表演并不总是为了社会身份认同，也可能是技术性、交际性的策略选择。

综上所述，前化变体作为一种语音现象，其意义早已超出纯语言学范畴。它可以是性别表演的产物，也可以是社会规范的引述行为，抑或是网络文化中的传播实践，甚至可能是个体在声音环境中作出的技术性选择。在 Butler 的框架中，我们得以理解其为何能如此广泛地激起讨论——它触碰的是性别的构造机制，是“女性如何说话才像个女性”的文化神话。

也正是在这种重复中， $t\epsilon$  组声母前化变体既体现了主流性别范式，也在某些偏差中折射出身份表演的特性。是否“做作”、是否“可爱”、是否“尴尬”，都反映了我们如何在听觉层面认同或拒斥某种性别建构。它既是传统规训的回应，也是主体调动语言资源、协商自我呈现的策略性实践。

### 3.5 “男国音”：北京青年 $t\epsilon$ 组声母的后化现象的初步观察

在调查  $t\epsilon$  组声母的前化现象过程中，我们意外的观察到一种与之相对的发音现象，即  $t\epsilon$  组声母的后化变体。我们暂且命名为“男国音”。具体而言，该现象表现为  $t\epsilon$  组声母出现后化的倾向，由齿龈音转为硬腭音，听感接近 [c $\epsilon$ -]、[c<sup>h</sup> $\epsilon$ -]、[ $\epsilon$ -]。尽管本研究的原始设计并未将此现象纳入调查范围，且因此未能系统性地记录其发生频率与年龄分布，但有两个值得注意的观察结果：

1. 在 44 名参与者及 10 名受访者中，此类发音仅在男性说话人中出现，女性说话人未见类似现象。

2. 使用“男国音”的说话人往往在语音上表现出对传统标准音的松动，且可能与性别身份、社会认同等因素有关。

这一现象是否构成系统性的语言变体，是否具有社会语言学意义，以及它是否与“女国音”形成某种性别化对立结构，尚需更多数据支持与深入研究。我们在此仅将这一观察作为研究的附注记录，期待未来研究能进一步追踪和分析“男国音”的性质与成因。

## 四 总结讨论

### 4.1 研究问题的回顾与回应

本文围绕两个核心问题展开：其一，女国音是否依然展现出显著的性别偏向？其二，传统解释框架（如“爱美心理”“方言迁移”或“京剧残留”）在今日语境下是否仍具解释力？通过10位初高中学生的深度访谈与声学分析，本文试图在社会语言学与性别理论的交汇点上，重新理解这一特殊变体的语言位置与社会意义。

首先，在性别分布方面，虽然女国音仍主要出现在女性语音中（57.1%），但男性使用者（26.1%）的出现打破了1980年代男性为零的格局（胡明扬，1987）。这一变化不仅预示着语言与性别之间关系的重新配置，也提示我们——性别并非语音变体的天然边界。男生对这一“女性化”语音的使用或是出于戏仿、或是无意识模仿，其动机的模糊性反而强化了Judith Butler所提出的“性别表演”的张力：当男性在言语中挪用传统“女性发音”，性别之为表演的本质便暴露出来。tɕ组声母前化不再是某种稳定性别身份的标志，而是一种被实践、被挪用、或被误读的风格资源。

其次，在动因方面，研究结果对传统解释路径提出了质疑。“爱美心理”虽在部分女性受访者中得到一定的呼应，但对于男生而言，这一解释力被显著削弱。他们更常提到的是模仿、社交适应或省力发音等动因。另一方面，方言影响与京剧残留理论也难以成立——多数使用者来自不区分尖团音的华北、东北地区，他们自身也并不具备南方语言背景，而现在学生也很少看京剧这种几乎已经脱离青年群众的艺术了（于海阔，2013）。尽管部分传统研究认为tɕ组声母前化变体可能源自老国音或京派语音（曹耘，1987），即民国时期官方推行但尚未全国统一的发音规范，但从当前语音传播路径与代际间语感差异来看，这一解释或难以成立。第一，作为语音输入源，老国音相关的影视剧或广播语言在当前青少年受众中几乎已不具可感知度，亦未在受访者的语用回忆中出现。第二，从语言社会化角度看，当代青少年的语言模仿对象更偏向社交平台上的同龄群体与网红话语，而非跨代的权威广播腔或京派朗诵风（Zhang, 2005）。因此，将tɕ组声母前化变体现象归因于老国音遗存，不仅缺乏直接的感知证据，也忽视了语音时尚的时代建构机制。这些证据表明，女国音的扩散路径已从“语言变异的自然迁移”转向“社交风格的主动调度”，其成为一种可被识别、模仿、再生产的“符号”。

最后，本文在分析过程中也拓展出数个新问题：多数女国音使用者未能意识到自身的发音差异，提示这类变体或许具有“前意识性”特征；其次，受众对其态度分化（可爱 vs. 做作），表明其正游走于标准化与污名化的边缘；而男生的使用行为，则进一步冲击了声音性别一一对应的传统预设。这些发现不仅回应了原有研究问题，也推动我们对语言与社会互动机制的更深层理解。

## 4.2 主要发现

在深入访谈与质性分析的基础上，本文归纳出如下关键发现：

### 1. 性别边界的模糊化

女国音虽依然以女性使用为主，但男性使用者的显著出现，预示着语言性别化格局的松动。男生对于女国音的使用往往出于模仿或调侃，其行为本身既可能是对性别规训的挑战，也可能是对“可爱”风格的再性别化演绎。这种行为既呈现出语言风格的去性别化，也暴露了性别规范的表演性基础。

### 2. 传统解释路径的失效

爱美心理难以解释男性使用；南方方言与老国音也与使用者实际语言背景不符。本文认为，这些解释忽视了当前社交媒体语境中，语言风格的流动性与模因化特征，无法捕捉前化变体作为“社交符号”而非“语言遗留”的本质。

### 3. 社交平台作为传播机制

不同用户群体在社交媒体平台的偏好（小红书/微博 vs. B站/游戏社区）反映出女国音的使用与某种审美倾向、语言风格之间的深度耦合。模仿对象（如台湾明星或网红郭老师）所体现的，是前化变体在可爱、精致与社交的维度中被构造、被流通的过程。

### 4. 声音的“功能性动因”

受访者教师J提出/s/音在嘈杂环境中更抗噪、发音更省力，这一观点为前化变体的使用提供了除性别外的全新视角，也与刘雅琦等(2024)关于/s/-/ʃ/声学性能差异的研究相互印证。它提醒我们语言选择并不总是审美驱动，也可能是感知与环境调节的功能性结果。

### 5. 无意识使用与语言意识的断裂

超过半数的使用者难以察觉自己发音到底属于哪一类。这反映出女国音可能已经渗入大众对于标注普通话的认知。语言变异具有前意识性，说明这种变异并非刻意模仿，而是通过社群交往在无意识的扩散。

### 6. 态度分化与变体定位的模糊

虽然部分受访者认为其“可爱”“精致”，但也有人认为其“做作”“不自然”。整体来看，女国音尚未成为稳定的“社会方言”或“语言风格”，而仍处于流动、竞争与重构的过程中。

## 4.3 理论与实践意义

本研究最核心的理论贡献，在于对 Judith Butler 的性别表演理论的语音层面实践验证。传统性别语言研究常将发音差异视作性别身份的语言体现，而 Butler 视角下的“性别”并非固定身份，而是一种通过日常行为不断重复构建的表演性机制(Butler, 1999)。在本研究中，男生对女国音的使用（无论是有意模仿、戏仿，还是出于审美或省力动因），都打破了“女性发音=女性身份”的预设，将性别

从语言形式中解放出来，使声音本身成为性别实践的场域而非标识。这不仅对传统性别范畴的划界提出挑战，也提示我们重新思考语言在性别建构中的多元角色。

同时，本研究也对语言意识与语言内隐性理论作出了回应。超过一半的前化变体使用者并未意识到自身的发音选择，这一现象说明：语言变体并不总以“风格选项”的形式被操控，而可能在社群互动与模仿机制中以“前意识”的方式流布。这为理解变体扩散的“非策略性路径”提供了语料支持，并可与语言习得研究中的“无意识输入”机制建立联系。

本研究在既有“女国音”研究的基础上实现了多个层面的理论补充与方法论突破。相较于20世纪80年代胡明扬(1987)、曹耘(1987)等研究将前化变体视为历史残留，本研究尝试将其重新纳入当代语言实践之中，从社会语言学、语言意识、性别表演与媒介传播等多重视角加以剖析，不再将其视作语言系统中的边缘噪音，而是作为一种结构性存在加以对待。尤其在性别维度上，研究对象从女性的发音特征转为性别化语言实践的动态过程，关注前化变体如何成为性别身份的操演媒介、如何被不同性别的使用者挪用、戏仿乃至重构，从而突破了以往将语音与性别机械对应的局面。

此外，研究也拓展了对变体传播机制的理解，指出新媒体平台（如小红书、微博）对语言风格的偏好正在重塑青年语音实践，可爱、软萌等标签背后的符号政治成为前化变体扩散的重要驱动力。这一发现促使我们重新思考语言传播的动因不再是自上而下的语言政策或教育系统，而是平台算法、社交圈层与媒介风格的合力作用。同时，研究揭示了语言使用中的意识层级问题——超过半数使用者并不知晓自身的发音偏差，这与语言习得过程中的输入机制及语音觉知能力有关，也提醒我们未来对变体研究应更多纳入使用者的主观经验与认知反馈。

在实践层面，研究的发现对语言教育、性别观念及语言规划均具有潜在影响。语音教学长期聚焦于标准音的灌输，却忽视了对学生语音自觉能力的培养，本研究所揭示的前化使用的无意识性说明音系觉知应成为未来普通话教育的重要组成部分。与此同时，男生使用前化变体的现象也对性别教育提出了新的挑战。在现有教育体系中，语言行为往往被纳入性别规范的评价体系，前化变体若被误读为“女性化”或缺乏“阳刚”，可能引发对学生性别表达的误判与规训。本研究主张建立更加宽容与多元的性别语言观，承认语言风格本身的开放性与可操演性。再者，前化变体与特定社交平台、美学趣味的绑定，也提醒我们关注社交媒体在塑造语言规范中的结构性角色，尤其是在青少年语言生态中，其影响甚至可能超过传统语言权威。尽管本研究未系统探讨社会阶层与语言选择之间的关联，但观察到前化使用者多具备较强媒介素养与都市中产特征，这一趋势提示前化变体有可能正在成为一种新的社会风格标记，值得在后续研究中深入追踪其在社会分层、群体认同建构中的作用。

#### 4.4 研究局限

尽管本研究尝试从多角度描绘  $t\phi$  组声母前化变体在当代青少年中的使用格局与社会动因，并辅以系统的访谈与执行分析，但仍存在若干限制，需在此诚实呈现并供后续研究反思。

首先，样本构成受限于研究者的实际调研能力，尽管覆盖了44位学生和10位受访者，性别与使用特征基本均衡，但研究对象集中于北京地区两所中学的初高中学生，地域与社会阶层的代表性有限，

难以囊括更广泛的社会语言差异，亦难推断该语音变体在全国范围内的传播路径与社会意义。对 44 位学生录音时由于环境限制，录音背景嘈杂，有可能影响听感判断。我们也未能向学生们示范应有的语速和词间停顿时常，导致不同学生朗读语速虽整体大致相似，但个别学生朗读语速略快，可能对前化发音有一定的影响。

受访者的语言使用主要基于研究者主观的知觉判断与访谈自述，尽管有一定交叉验证，但未能引入声学实验等技术手段对语音进行具体声学参数的量化分析。这可能在一定程度上限制了对发音细节与使用频率的客观描绘，尤其在模糊使用者与无意识使用群体中，音变的边界尚不清晰。

此外，我们的访谈时间均控制在 10 分钟左右，可能限制了对个体语言经历、成长背景与媒介接触史的深入探讨，导致部分语言行为的解释仍显悬浮，未能构建出更扎实的语言社会化链条。

最后，由于研究聚焦于普通话体系内部的细微变体，未能充分纳入其他语种或方言对该现象的可能影响，也未对  $t\epsilon$  组声母前化变体是否在非汉语族群或多语言环境中呈现类似特征进行考察，限制了其跨语言的普遍性或独特性判断。理论使用上，尽管本研究尝试引入性别表演理论与新媒体语言观念，但尚未系统整合语言经济学、语言资本理论等框架，以更全面地理解变体背后的社会结构驱动与价值再生产机制，这亦是值得未来研究补充之处。

## 4.5 未来研究方向

基于上述局限与当前研究发现，未来关于  $t\epsilon$  组声母前化现象的研究可在三个维度上进一步推进：其一，在样本拓展层面，应超越北京单一城市、初高中学生的受限框架，采集更广泛的年龄群体与地理区域，尤其是成年女性与职场语言使用者，以探讨前化变体是否在更大范围内形成世代、性别或行业性的语言风格标记。与此同时，也应关注更多未被充分描写的男性使用者，厘清其语言行为背后复杂的性别意识与身份建构策略。

其二，在研究方法层面，应鼓励语音实验方法与社会语言学方法的深度结合。一方面可引入语音声学分析、感知实验、自动语音识别等技术手段，对发音的物理属性、清晰度与感知界限进行精确测量，提升研究的客观性与可重复性；另一方面，应继续强化访谈与田野观察中的细节采集，构建语言实践与社会情境之间更有力的关联链条，尤其是在语言风格、性别表演与媒介消费的交汇处深入挖掘，避免将变体使用简化为静态的个体偏好。

其三，在理论建构层面，未来研究可尝试将  $t\epsilon$  组声母前化变体置于更宏观的语言变异理论之中，作为一种正在浮现的“中产都市语言风格”进行理解，与饭圈用语等网络时代新兴语言变体共同构建出数字语境中的社会语言谱系。此外，可进一步引入布迪厄的语言市场理论、Silverstein 的“语言意识形态”框架，思考语言变体如何在社会价值体系中被赋予符号资本，如何在性别与社会地位的交叉结构中被使用、评价、流通与再生产。最终，这些延展性研究将有助于我们不仅理解“女国音”作为一个音变现象本身，更理解它在一个急剧变化的社会语境中所扮演的角色与承担的张力。

最后，在本研究过程中，我们还初步观察到一种与“女国音”相对的发音现象，我们先暂称为“男国音”。尽管目前相关资料有限，无法形成系统性分析，但该现象的存在提示我们，性别差异对标准

音变异的影响可能不仅体现在前化这一方向上，也可能呈现出更复杂的、甚至互为镜像的音变模式。未来研究可以在以下几个方向展开：其一，系统收集“男国音”现象的语料，明确其发音特征、分布范围与频率；其二，探讨该现象的社会动因，如性别身份建构、语言态度、地域背景等因素；其三，考察“男国音”与“女国音”是否构成一组互补或对立的社会语言学变量。我们相信，对“男国音”的进一步研究，将有助于深化我们对性别与语言标准之间关系的理解。



## 致谢

首先，我们衷心感谢来自北京市第八十中学和北京市朝阳区外国语学校，参与本研究的 44 名同学，你们在数据采集过程中给予的支持和配合，为本研究提供了宝贵的语音样本和宝贵的第一手资料。你们的认真和热情让这项研究得以顺利开展。

其次，我们要特别感谢在数据采集过程中给予我们帮助的同学和老师。你们不仅在参与过程中为我们提供了宝贵的建议，还在研究方法、数据收集等方面为我们提供了实质性的支持。尤其是在繁忙的学业中，你们依然抽出时间参与，让我们深感感激。

特别感谢在访谈过程中提供宝贵意见的 10 位同学和老师。你们的真诚分享和深入思考使我们能够更清楚地理解 *tc* 组声母前化现象背后的社会文化动因，给我们的研究带来了不可替代的启发和视角。

此外，我们也要向曾经在学术上给予我们启发的前辈们表示感谢。你们的研究成果为我们提供了宝贵的理论参考，你们严谨的学术态度也时刻激励着我们不断追求学术深度和严谨性。正是有了你们的前行铺路，我们才能够在这条道路上继续前行。

再次感谢所有参与者与支持者，正是因为你们的帮助与参与，这项研究才得以顺利完成。

## 参考文献

- Androutsopoulos, J. (2014). *Mediatization and Sociolinguistic Change*.
- Bloor, M., & Wood, F. (2006). *Keywords in Qualitative Methods: A Vocabulary of Research Concepts*. SAGE Publications Ltd. <https://doi.org/10.4135/9781849209403>
- Butler, J. (1999). *Gender Trouble: Feminism and the Subversion of Identity* (10th Anniversary Edition, 2nd). Routledge. <https://doi.org/10.4324/9780203902752>
- Coupland, N. (2001). *Language, situation, and the relational self: Theorizing dialect-style in sociolinguistics*. Cambridge University Press.
- Eckert, P., & McConnell-Ginet, S. (2013). *Language and Gender* (2 本). Cambridge University Press.
- Krippendorff, K. (2019). *Content Analysis*. SAGE Publications, Inc. <https://doi.org/10.4135/9781071878781>
- Labov, W. (1972). *Sociolinguistic Patterns* (6th printing). University of Pennsylvania Press. <https://archive.org/details/sociolinguisticp00will>
- Labov, W. (1982). *The Social Stratification of English in New York City* (1st 本). The Center for Applied Linguistics. <https://doi.org/https://doi.org/10.1017/CBO9780511618208>
- Ladefoged, P., & Johnson, K. (2015). *A Course in Phonetics* (7th 本). Cengage Learning.
- Lee, C. (2016). *Multilingualism Online* (1st 本). Routledge. <https://doi.org/10.4324/9781315707211>
- Liu, M. (2022). Sociolinguistics in China. 收入 *Oxford Research Encyclopedia of Linguistics: Oxford Research Encyclopedia of Linguistics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.914>
- Milroy, L. (2002). *Authority in Language: Investigating Standard English* (3rd 本, 页 xii+208). Routledge. <https://doi.org/10.4324/9780203463255>
- Podesva, R. J. (2011). The California Vowel Shift and gay identity. *American Speech*, 86(1), 32–51. <https://doi.org/10.1215/00031283-1503741>
- Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona 1. *Journal of sociolinguistics*, 11(4), 478–504.
- Zhang, Q. (2005). A Chinese Yuppie in Beijing: Phonological Variation and the Construction of a New Professional Identity. *Language in Society*, 34(3), 431–466. <https://doi.org/10.1017/S0047404505050153>
- 中国社会科学院语言研究所, 中国社会科学院民族学与人类学研究所, & 香港城市大学语言资讯科学研究中心 (编). (2012). 中国语言地图集 (第 2 版). 商务印书馆.
- 于海阔. (2013). 地方戏曲是否应固守方言. 中州学刊, 196(4), 161–166.

- 刘雅琦, 王韞佳, & 余俊毅. (2024). 声学特征与普通话辅音声母的感知混淆. 声学学报, 49(6), 1-13.  
<https://doi.org/10.12395/0371-0025.2024128>
- 张振兴. (2013). 方言研究与社会应用. 商务印书馆.
- 徐世荣. (1956). 普通话语音基本知识纲要(二). 语文学习, 9, 14-16.
- 徐世荣. (1979). 普通话语音和北京土音的界限. 语言教学与研究, 1, 7-22.
- 曹耘. (1986). 北京话语音里的性别差异. 汉语学习, 6.
- 曹耘. (1987). 北京话 *tc* 组声母的前化现象. 语言教学与研究, 3, 8.
- 朱晓农. (2004). 亲密与高调——对小称调、女国音、美眉等语言现象的生物学解释. 当代语言学, 3, 193-222.
- 胡明扬. (1987). 北京话“女国音”调查(1987). 语文建设, 1. <https://doi.org/10.16412/j.cnki.1001-8476.1988.01.014>
- 陈松岑. (1985). 社会语言学导论. 社会语言学导论.
- 高本汉, 赵元任(译), 罗常培(译), & 李方桂(译). (1940). 中国音韵学研究 (第一版). 商务印书馆.

## 附录

### 附录 A：访谈文本（节选）

受访者 T（男，学生，17 岁，标准发音）

采：那么，我们第一个问题是：你有没有注意到你身边人——或是在网上看到的，任何地方——有没有人把在读机器人、骑车、西红柿这三个词的时候读起来像机器（/tsi tsh/）人、汽（/tsh/）车和西（/si/）红柿？

T：OK，我觉得有一些人确实会读成这种音。他们可能想把每一个字都发的特别的慢，特别的清楚，然后就会有这种的效果，可能会拖长声音这种。

采：那你第一次听到这种类似的发音是在哪里呢？比如说，是在身边人口中听到的，是短视频只看到的，还是在什么其他的地方？

T：嗯，可能是在短视频，确实短视频有一些。有些人他们可能会过分的做作，就是他们可能拖得更长，所以声音会更明显一些。

采：OK。所以你感觉这种发音习惯更多是做作，而不是就是像类似可爱这种。

T：可能有些人会把它认为可爱，我觉得我没有任何贬义的意思，就是可能就是他比较想强调这个音，然后他可能觉得这个音比较要更让别人舒服一些，或者是更能引起注意。

采：明白。那这种发音你会认为它是奇怪的吗？还是就是一种比较普通的现象？

T：我觉得比较普通就是，对，就是我，我可能知道他为什么要这样，所以我就觉得他比较普通。

采：好的。那我们进行下一部分。你认为你自己说话的时候会不会有这样的发音？

T：我自己说话是吗？中文可能没有，但是英文有时候会有。但是不是这种类型的，可能会特意模仿在某些单词的时候可能会特意模仿某种语调，来凸显出我的 legitimacy，显得我比较 native，所以肯定是有。

采：哈哈，OK。那么你认为这种说话方式它是如何传播的？或者说，它是从哪里起源的？

T：我觉得主要是网红，就是可能就有些网红他们说话就特别怪，就比如装腔作势，然后会有一些人就去 copy that，就会去学那个声音。有可能只是玩梗，也可能就是别人就是那样说的。

采：是的，因为他们的影响力也比较大。

T：对。

受访者 Z（男，学生，16 岁，标准发音）

受访者 X（女，学生，17 岁，完全前化）

采：你们有没有注意到你或者是你身边的人说话，有时候会把机器人、骑车、西红柿这样的词读成机器（/tsi tsh/）人，汽（/tsh/）车，和西（/si/）红柿这样的现象？

Z: 我, 我有[注意到这样的现象]。我初中同学有。

X: 没有吧。

Z: 但是这不就是很常见的吗?

采: 那你第一次听到这种发音方式, 是在你初中同学的口中, 还是什么短视频啊、是身边人是在哪里听到这种? 最开始第一次听到这种。

Z: 我估计是初中同学。我初中也没有什么社交, 然后就只[注意到了]初中同学[这么]说。

采: 你呢?

X: 我好像没有[注意到有]人这么说话。

采: 没有注意到有人这么说话是吗? OK。那么下面这个问题就只针对受访者 Z。你认为这种说法它是从什么地方起源的? 它是怎么传播起来的? 是通过网红还是通过朋友这种面对面的交流这样?

Z: 有没有一种口音就是这么说话的?

(采访 Z 和 X 时, T 也在旁边)

T: 我突然想我能加一句, 我刚想感觉就是台湾有些口音。

Z: 对对对, 我其实也想[说]台湾。

采: 那就是方言的影响。

Z: 对, 然后有些人可能觉得这么说话很好听.....

X: 对, 比较可爱, 就会改成这样。就是网上抖音很火的, 会学他们说话, 那就是从网红这边传过来的。

采: OK, 所以就是像模仿那种台湾电视台, 译制腔的那种。

Z: (对 X) 你能模仿一个吗?

X: 我之前会说台湾话, 但是我现在不会了, 因为一直之前一直追星来着。啊? 不会是我像这个吧?

Z: (笑) 说着说着就说出来了。

X: 你看我没有像啊, 我觉得还好。

采: 那你们对于这种方言它的态度是什么样的?

X: 可爱。因为我追星追台湾人, 我觉得他们说话很好听啊。

Z: 你[追]哪个星? 台湾人?

X: (某台湾明星) 呀。

Z: 哦哦, 对对对, 我忘了。

采: (对 Z) 你觉得呢?

Z: 我觉得还好, 不能说可爱, 但也没有很做作, 就这么说话就正常。

受访者 M (男, 学生, 16 岁, 后化发音, 即“男国音”)

受访者 T (女, 学生, 16 岁, 半前化)

受访者 W (男, 学生, 17 岁, 标准发音)

受访者 S (女, 学生, 16 岁, 标准发音)

采: 你们第一次听到这种的发音是在什么地方?

T: 我感觉像就歌里会有的。就你听着听着, 然后你的发音会往那边[靠]。然后有的人提出来的时候, 你就会发现好像网上很多歌手都这么唱。

采: 你们认为你们自己说话的时候会不会有带有这样的发音?

W、S: 不会。

T: 我觉得我说话没有, 那你要说真的到歌里头, 我觉得可能如果说真听多的还是会有的, 我觉得这个对, 就是受你身边影响会很大, 就如果你真的常听这个, 他确实会[对你有影响]。

采: 你们认为这样的说话习惯是怎么传播起来的?

S: 很多都是人传人的吧。

T: 因为我在身边没有见过, 所以我不知道他对身边的影响会有多大, 但如果说真的是, 我觉得他至少从网上来说是有影响的, 然后我觉得他可能其实更多偏向于地区类的, 就我觉得他可能是一个地区口音的[影响]。我觉得他有可能其实是从某个口音出来的, 我的猜测啊。然后可能是某一类人会[这么]说话, 就相当于口音是有“传染性”的, 像东北话的“传染性”。

S: 诶? 东北话怎么了?

T: 就是传染性很强, 我没有批判。

W: 东北话有啥传染性?

T: 就是说, 如果你有一个东北室友, 你可能就会被带着说东北话。

采: 那你们觉得这种发音, 你们对于这种发音的态度是什么?

M: 做作的。

T: 我觉得有点做作。

W: 反正肯定不是可爱的, 我觉得更多是习惯上的。

采: 好的。非常感谢你们的分享, 这对我的研究非常有帮助。

受访者 L (男, 学生, 17 岁, 半前化)

受访者 O (女, 学生, 17 岁, 标准发音)



采：非常感谢你们参与这次访谈。我正在做一个关于北京年轻人说话习惯的小研究。这个录音不会公布你们的真实姓名，只用于学术研究。首先想问，你们有没有注意到，自己或者身边的人会把“机器人”“骑车”“西红柿”这些词发成更尖、更促的音，比如“机器（/tsi tʂʰi/）人”、“汽（/tʂʰi/）车”和“西（/si/）红柿”这样的？

L：（模仿台湾口音）我们一般不叫西（/si/）红柿（/ʂʰi/）诶，我们都叫番茄（/tsjɛ/）（笑）。

采：（笑）那所以说你们有没有注意到有人会这么说？

O：（某同学）。

L：真的吗？但是你这么一说，好像一下想不起来具体例子，但我觉得应该是有的。

采：对对对，那你们有没有注意到身边或者网络上有谁发音特别尖，听起来挺明显的？

O：我觉得更多时候是在大家一起聊天，说到一个搞笑或有趣的事情时，他们会刻意把词说得特别一点，可能是为了制造共鸣或搞笑效果。网络影响也挺大的。其实并不是真的说话习惯，更像是一种表达方式。

采：所以更多是为了强调或搞笑，不是生活中的自然用法。那你们第一次接触这种说法，是通过身边人还是在短视频或影视剧里听到的？

O：我是短视频。

L：我不记得……我。

采：所以不是特别明显的一个来源，对吧？那你们自己说话时，会不会也这样发音？

O：我自己平常不会这样说。但如果是在聊天，朋友这么说了，我可能会下意识模仿，一方面是为了拉近关系。

采：噢，appealing to pathos。你呢？

L：有时候可能嘴瓢，舌头捋不过来，也不是故意的。比如说（演示嘴瓢）。

O：对，也有短视频的影响。有时说话时脑子里突然浮现出那个视频的画面或者声音，就模仿了。

采：所以你们觉得，这种说法主要是短视频或网红文化带起来的，对吗？

O：可能我刷到的视频比较“低质”，但的确是我第一次接触到这种说法是在短视频里。印象最深的也来自短视频。日常偶尔会听到朋友模仿，但比较少。

采：OK，那你们对这种说话方式的态度是怎样的？觉得可爱、做作、土，还是有点奇怪？

L：我觉得这只是个说话方式，没必要上纲上线。但如果有人为了营造某种“人设”而故意这么说，比如为了显得“精致”，那我觉得就有点哗众取宠。我个人不太喜欢那种做作的感觉。不过图一乐开个玩笑也没什么。如果是无意识的，或者受短视频影响的一时模仿，我觉得还挺可爱的。

采：嗯，你也是这样认为的吗？

O: 差不多。如果它成了一个网络梗,比如“京爷”(夸张模仿北京口音)、“国际生说话”所谓的那类,大家模仿一阵我觉得没问题,挺正常的。但如果有人真的把这种说话方式当作习惯,甚至故意“打造人设”,那就有点奇怪了,跟这样的人相处也会有点不自在。

采: 好的,录音有了,后续我可以再回听确认。谢谢你们的分享,这对我的研究很有帮助!

L: 别忘了写 acknowledgment!

采: (对 L) 其实你说你自己没有,但你偶尔说话的方式也会带一点[前化]。

O: 比如什么时候?

采: 刚刚就有。回头我给你调录音听(笑)。今天就到这儿,谢谢你们!

### (结束后补录)

采: 没事没事,你直接说,你就说就行了。

O: 就是我最早的印象是——我不知道这俩是不是一回事儿——就是会把字给它变个调,或者变个音。比如说把“妈妈”变成“妈沫”这种,然后就是每个词基本上都稍微给它变一下。我不知道这不是有方言的因素在,可能有一点,但我觉得更多还是当时的环境影响。

采: 你说的是大概哪段时间?

O: 疫情那段时间。因为大家都很无聊,都在网上刷手机。那时候有一个叫郭老师的人啊.....

L: 对对对!我不太[喜欢她]。

采: 但她对这种说话方式这几年的传播,影响真的特别大。当时她正好最火,我初中好多女生都学她那种说话方式。

O: 她开直播,每条视频都有那种独特的风格,就感觉已经创造了一个属于她自己的语言了。所以我觉得短视频的效应真的挺强的,而且人们会不自觉去模仿。不管是有意模仿那个网红,还是在日常生活中说话时,模仿成了习惯。

O: 对,你会发现你身边大部分人其实都受到了这种影响。那么,如果你想融入这个群体,有时候即使你内心不太情愿,可能也会说一些,或者做一些类似的事情。所以在某种程度上,我们其实都在变相地推动了这种传播。

### 受访者 J (女, 教师, 30 岁, 完全前化)

采: 好的,那我们今天就以聊天的方式进行,不用紧张。今天这个小研究是关于北京年轻女性的说话习惯,不是普通话考试,也不会公布录音或个人信息,仅用于学术研究。没问题吧?

J: 没问题。

采: 好,那我们开始吧。第一个问题是:你有没有注意到身边或网上的一些人,会把像机器人、骑车、西红柿这样的音发成像机器(/tsi tʃi/)人、汽(/tʃi/)车、西(/si/)红柿的发音?

J: 有啊, 其实我就是。但是如果你要我认真说的话, 我当然会说“机器人”。但如果我累了, 比如讲课讲得很累, 就可能说成机器 (/tsi tʃi/) 人、汽 (/tʃi/) 车, 会发成一个 /si/ 的音。

采: 嗯, 我之前访谈的一些人说他们只有在和朋友聊天、想要强调或者制造幽默的时候才会用这种发音。你用的时候.....

J: 为了省事吧, 也算是一种放松。你要发 /ci/, 舌头得动得多, 用 /si/ 就省劲儿多了。

采: 你第一次听到这种发音, 是从哪儿来的? 是身边的人、短视频、还是电视剧?

J: 这得从我小学说起了。当时我们班有个学习课代表就是这种发音。我那会儿其实不太喜欢这个发音, 但后来发现它真的挺省事的。

采: 所以也算是被潜移默化地影响了?

J: 也没那么潜移默化啦, 其实我是最近才开始用这个音。有时候懒得动嘴, 这种发音是最好的选择。

采: 那你身边除了你以外的人, 也会这样说话吗?

J: 很少。我身边的人大多是老师、律师, 还有一些做短视频的媒体人。他们说话比较讲究, 要么职业需要, 要么想让人听清楚, 所以不会省这个力。

采: 那学生呢?

J: 女生里会有一些, 尤其是比较在意自己形象的那种。

采: 在年级分布上有区别吗?

J: 高年级更多一些。低年级的话, 老师会比较严格地纠正发音, 比如一年级、二年级的老师绝对不会让他们这么说话。但像你们这个年龄段就可能会多一点。我记得(某同学)就有点这种发音, 我不确定, 但感觉她好像有一点。这种懒懒的、但又好看的气质很符合她的形象。

采: 你觉得这些女生是怎么学会这种发音的? 是从短视频、网红那儿学的, 还是朋友之间口口相传?

J: 我更 prefer 后者。短视频我没怎么注意过有这种现象。可能用这个音的话, 在设备里容易爆音, 尤其是录音质量差的时候。我录课的时候就 very 注意这个, 不会用这种发音。

采: 你个人怎么看这种发音方式? 你觉得它是因为省力呢? 还是觉得说这种话的人可爱、或者故作?

J: 我觉得都不是。我一般会觉得这个人可能是累了。我也不会特别在意, 也想不起来我身边谁有这个习惯。对我来说就是个无关紧要的点。

## 附录 B: 数据分析代码

由实际使用代码修剪而来，保留了核心功能，去除了不重要的功能，如适配字体、打印调试等。

```
import pandas as pd
import numpy as np
from scipy.stats import chi2_contingency, mannwhitneyu, ttest_ind
try:
    from statsmodels.stats.proportion import proportions_ztest
except ImportError:
    def proportions_ztest(counts, nobs):
        p1 = counts[0] / nobs[0]
        p2 = counts[1] / nobs[1]
        p_pool = (counts[0] + counts[1]) / (nobs[0] + nobs[1])
        se = np.sqrt(p_pool * (1 - p_pool) * (1/nobs[0] + 1/nobs[1]))
        z = (p1 - p2) / se if se > 0 else 0
        from scipy.stats import norm
        p_val = 2 * (1 - norm.cdf(abs(z)))
        return z, p_val
from collections import Counter

# == Read data ==
df = pd.read_csv('data.csv')

# == Data preprocessing ==
def clean_and_split(text):
    if pd.isna(text):
        return []
    items = [item.strip() for item in str(text).split(',') if item.strip()]
    return items

df['social_media_list'] = df['social_media'].apply(clean_and_split)
df['content_list'] = df['content'].apply(clean_and_split)

# == Counting function ==
def count_items_by_group(df, column, group_col='group'):
    result = {}
    for group in df[group_col].unique():
        group_data = df[df[group_col] == group][column]
        all_items = [item for sublist in group_data for item in sublist]
        result[group] = Counter(all_items)
    return result

# == 1. Social media usage analysis ==
social_media_counts = count_items_by_group(df, 'social_media_list')
all_platforms = set().union(*[counter.keys() for counter in
social_media_counts.values()])
group_a_total = len(df[df['group'] == 'A'])
group_b_total = len(df[df['group'] == 'B'])
```

```
df['social_media_count'] = df['social_media_list'].apply(len)

# === 2. Content preference analysis ===
content_counts = count_items_by_group(df, 'content_list')
all_contents = set().union(*[counter.keys() for counter in content_counts.values()])
df['content_count'] = df['content_list'].apply(len)

# === 4. Two-proportion z-test for each platform ===
platform_test_results = []
for platform in all_platforms:
    a_users = social_media_counts['A'].get(platform, 0)
    b_users = social_media_counts['B'].get(platform, 0)
    if a_users > 0 or b_users > 0:
        counts = np.array([a_users, b_users])
        nobs = np.array([group_a_total, group_b_total])
        z_stat, p_val = proportions_ztest(counts, nobs)
        platform_test_results.append({
            'platform': platform,
            'A usage rate (%)': round(a_users/group_a_total*100, 1),
            'B usage rate (%)': round(b_users/group_b_total*100, 1),
            'z-statistic': round(z_stat, 4),
            'p-value': round(p_val, 4),
            'Significant': 'Yes' if p_val < 0.1 else 'No'
        })

# === 5. Two-proportion z-test for each content type ===
content_test_results = []
for content in all_contents:
    a_users = content_counts['A'].get(content, 0)
    b_users = content_counts['B'].get(content, 0)
    if a_users > 0 or b_users > 0:
        counts = np.array([a_users, b_users])
        nobs = np.array([group_a_total, group_b_total])
        z_stat, p_val = proportions_ztest(counts, nobs)
        content_test_results.append({
            'content': content,
            'A viewing rate (%)': round(a_users/group_a_total*100, 1),
            'B viewing rate (%)': round(b_users/group_b_total*100, 1),
            'z-statistic': round(z_stat, 4),
            'p-value': round(p_val, 4),
            'Significant': 'Yes' if p_val < 0.1 else 'No'
        })

# === 6. Chi-square test of independence for each platform ===
chi_square_results = []
for platform in all_platforms:
    a_use = social_media_counts['A'].get(platform, 0)
```

```

a_not_use = group_a_total - a_use
b_use = social_media_counts['B'].get(platform, 0)
b_not_use = group_b_total - b_use
contingency_table = np.array([[a_use, a_not_use], [b_use, b_not_use]])
if np.all(contingency_table >= 1):
    chi2_stat, p_val, dof, expected = chi2_contingency(contingency_table)
    chi_square_results.append({
        'platform': platform,
        'Chi-square statistic': round(chi2_stat, 4),
        'p-value': round(p_val, 4),
        'Significant': 'Yes' if p_val < 0.1 else 'No'
    })

# == 7. Other statistical tests ==
# Mann-Whitney U test for influence scores
group_a_influence = df[df['group'] == 'A']['influence']
group_b_influence = df[df['group'] == 'B']['influence']
mw_statistic, mw_p_value = mannwhitneyu(group_a_influence, group_b_influence,
    alternative='two-sided')

# t-test for the number of social media platforms used
group_a_social = df[df['group'] == 'A']['social_media_count']
group_b_social = df[df['group'] == 'B']['social_media_count']
t_stat, t_p = ttest_ind(group_a_social, group_b_social)

# t-test for the number of content types viewed
group_a_content = df[df['group'] == 'A']['content_count']
group_b_content = df[df['group'] == 'B']['content_count']
t_stat2, t_p2 = ttest_ind(group_a_content, group_b_content)

```