

徐以中 杨亦鸣\*

# 汉语和日语听觉反馈声学特点及语言认知习得能力——基于噪声 (Lombard) 听觉反馈条件实验研究

**摘要** 元认知是对认知的认知，语言元认知能力体现为说话人对自身言语行为的自我监控，在心理语言学中属于言语听觉反馈现象。本文基于噪声 (Lombard) 条件的听觉反馈实验，考察了操汉语和日语者的言语发声特点及其所反映的认知心理机制。统计结果揭示：一、噪声对话语语速和音高的影响不显著但对音强的影响非常显著；这个结果也验证了赵元任先生提出的“汉语音高可以区别词义使得汉语在噪声传播中是一个优点”的观点；二、语义和语境因素对音强和音高变化的影响具有显著意义；三、汉日两种语言被试在语言元认知能力上既有普遍性又有差异性。总之，开展言语听觉反馈研究对于揭示语言认知和习得机制具有重要意义。

**关键词** 噪声 (Lombard) 听觉反馈 认知能力 语言习得

DOI 10.1515/caslar-2014-0013

## 1 听觉反馈的内涵及功能

在探讨语言认知和习得机制的过程中，一般倾向于把言语活动划分为言语发声和言语感知两个阶段。过去的研究多集中于其中的某一方面而较少关注两者之间的互动 (interaction) (Carol and Bruno 2005)。

---

\*通讯作者：杨亦鸣：江苏师范大学语言科学学院教授，博士生导师，主要研究领域为神经语言学和理论语言学。

电子邮件：yangym@jsnu.edu.cn

通讯地址：221009 江苏省徐州市江苏省语言科学与神经认知工程重点实验室

徐以中：博士，1975年生于江苏沭阳，南京航空航天大学外国语学院副教授，硕士生导师，研究兴趣主要为认知语言学、实验语音学和对比语言学。

电子邮件：CalebXuxu@163.com

通讯地址：211106 江苏省南京市江宁区胜太西路169号南京航空航天大学外国语学院

人类认知神经系统在日常言语交际中不仅时刻在识别和理解其他人的言语，同时也在感知和监控自己的言语 (Balconi 2008: 159)。这种对自身语音的感知和监控便是听觉反馈 (auditory feedback)<sup>①</sup>。可见，“听觉反馈”是指说话者在说话的同时包含对自己声音的监视和调控。听觉反馈现象反映了言语感知和发声之间的互动。赵元任 (Chao 1964) 指出，语言是一种社会行为，当一个人感知到他自己的说话方式在被别人感知的时候，他的说话内容和说话方式会迅速受到短暂或长久的影响。Carroll (1994: 207–211) 也指出，在说话时，为了保证说出自己想说的话和保证自己的说话方式，说话人通常会对自己的话语进行监控 (monitor)。这种自我监控 (self-monitoring) 体现在三个方面：首先，当意识到自己言语中的错误时说话者会自我停顿 (self-interruptions)；其次，说话者通常会说些插入语 (editing expressions)，比如“啊”、“对不起”、“我的意思是”等；最后，说话者会对言语进行自我修正 (self-repairs)。

听觉反馈能力对正常的言语发声起着非常重要的作用 (Burke 1969; Burnett and Larson 2002)，尽管该作用由于不容易被言者所觉察而导致经常被低估或忽视。比如人们打电话时如果突然遇到嘈杂的噪声，说话者通常会自动提高自己的音量 (loudness) (Jones and Munhall 2002)。这是言语听觉反馈的一种典型表现。儿童在学说话的时候，都是通过边听边说，逐渐学会了讲话。如果他根本听不见自己的发音，就无法判断说话效果，也就无法学会讲话。聋人由于听觉系统的障碍无法随时监听自己发出的声音，因而也就无法学会语言。所以会出现人们常说的“十聋九哑”现象。

听觉反馈对发音人自己的重要性很早就通过实验得到了证明 (Lee 1950)。这种实验叫延时听觉反馈实验 (Delayed Auditory Feedback: DAF)。实验可简单描述为：一个人在说话时听觉感知系统会在瞬间及时听到自己的声音，但是，当说话人的声音传到自己听觉系统的过程中发生延迟时，譬如延迟50毫秒、200毫秒、400毫秒等，此时说话人的话语就会出现类似口吃的不流利现象 (speech dysfluency)。迄今这个实验已成为语音听觉反馈研究的一个经典范式。

在神经基础层面，人类的发声涉及神经系统对声道和呼吸肌肉的计划和调控 (Carroll 1994: 211; Guenther 1995; Jones and Munhall 2005)。在包括言语活动的任何一项运动中，反馈控制要求在执行过程中不断进行控制和修正 (Wolpert et al. 1995; Wolpert and Miall 1996; Max et al. 2004)。来自功能性神经影像学技术 (PET和fMRI) 的研究证据表明，人类发声时，甚至在发声之前，不仅大脑的运动区域被激活，而且大脑的听觉皮层 (颞叶) 也被激活 (McGuire et al. 1996; Price et al. 1996; Wildgruber et al. 1996; Hirano et al. 1997; Gunji

---

① 听觉反馈的广义理解有两种：一种是交际中对其他交际者 (interlocutor) 的话语反馈，另一种是对自己声音的反馈。本文听觉反馈含义专指后一种。此外，在口腔中形成的语音到达听神经中枢有两种传导路径，一种是通过骨骼肌肉的传输，一种是通过空气的传输，前者叫 somatosensory feedback，后者叫 auditory feedback (Tremblay et al. 2003)。有时两种反馈都叫听觉反馈，本文的听觉反馈专指后一种 auditory feedback。

et al. 2000, 2001; Jones and Munhall 2005)。总之，言语发声包括对来自听觉、运动、小脑 (cerebellum)、体觉 (somatosensory) 等信息的整合，但迄今为止，我们对控制言语发声的具体的感觉运动表征机制 (sensorimotor representations) 却所知不多 (Jones and Munhall 2005; Guenther 2006)。尤其是涉及两种语言被试的对比研究更加鲜见。

以往的一些研究虽然在一定程度上揭示了言语感知-运动的一些规律，但是人脑是个相当复杂的系统，在言语产生的过程中，反馈本质是什么，以及反馈与感知之间是如何配合完成任务的？学者们在这些问题上并没有形成一致意见 (Max et al. 2004)。在日本及欧美一些著名大学或研究机构有专门从事语音听觉反馈研究的实验室，如美国的西北大学、日本国际电气通信基础技术研究所 (ATR)，而中国国内的相关研究机构在这方面显然重视不够。

听觉反馈能力也是一种语言元认知能力。“元认知 (metacognition)”概念较早由 Flavell 提出。简言之，元认知就是对认知的认知，其实质是个体对当前认知活动的认知调节 (Flavell 1979, 1985; Sternberg 1994; 汪玲、郭德俊 2000)。相对一般认知能力而言，元认知能力属于一种自上而下的加工。在人的智力活动中，元认知处于支配地位，是智力的本质和核心。在语言认知中，语音感知作为一种认知活动，对它的认知便形成了对言语活动的元认知。在言语交际中体现为说话人对自身言语行为的自我监控。

对听觉反馈现象进行专门研究是从认知心理学界开始的。早期人们对听觉反馈感兴趣始于延时听觉反馈 (DAF) 实验 (Lee, 1950; Yates, 1963)。当初这项实验并未引起许多人的注意和重视，直到不久后研究者发现，如果对口吃者听觉反馈信号进行延迟，则可以有效地改善口吃发生程度 (Burke 1969; Kalinowski et al. 1993)。近期市场上出现可以有效降低口吃现象的产品“思比易 (SpeechEasy) 就是利用听觉反馈原理进行工作的 (Muellerleile 1981; Lincoln 2006; Packman 2012)。由于听觉反馈有了广阔的应用前景，因此引起了研究者极大的兴趣。此后，研究者发现，如果把说话人的言语基频提高或降低，然后让其边听边说，结果也会对言语发声带来影响 (Burnett et al., 1998; Jones & Munhall, 2000)。根据实验中给予被试反馈音刺激的不同特性，目前听觉反馈研究已形成三种实验范式，分别是延时听觉反馈实验、变频听觉反馈实验和噪声听觉反馈实验 (Junqua 1996; 河原英纪 2003)。

尽管以往的研究已取得不小成就，然而，目前人们对言语发声过程中听觉反馈的具体机制尚不明了 (Max et al. 2004)。而且，以往的研究都是基于认知心理学的视角进行的，真正从语言学角度切入的研究相当少见。值得注意的是，赵元任很早就注意到了听觉反馈现象并早在 1964 年就有一篇专门描述听觉反馈的论文，文章主要讨论“通讯技术对说话方式的反馈效应” (Some Feedback Effects of Communication Technology on Styles of Speech)。论文内容涉及在不同的条件下，话语的辅音、元音、基频等参数变化以及在噪声中说话方式的变化等 (Chao 2006: 784-791)。不过该论文总体上属于现象描述和理论探讨性论文，未能从实验的角度探讨听觉反馈的特点和机制。尽管赵元任对听觉反馈研究进行了开创性介绍和尝试，可是此后在语言学界仍较少得到研究者的

关注。本文接下来将专门基于噪声听觉反馈实验来考察言语声学特点及其所体现的理论价值。

## 2 噪声听觉反馈研究

语音在传输、采集的过程中，经常会受到噪声的干扰，由于噪声的干扰，许多语音处理系统的性能会严重下降（田斌、易克初 2003）。噪声也是日常生活中影响言语识别的最常见因素。过去在这方面的研究也比较多，但到目前为止噪声环境中的语音识别仍然是亟待克服的难点。

生活中各种各样的噪声不仅对识别他人的声音有影响，而且它通过影响说话人对自己声音的感知从而影响说话人的言语产出 (speech production)<sup>②</sup>。在交际中，当环境噪声加大时，说话人由于受生理和心理的影响，其语音声学特征会发生变异，如音强 (intensity) 加大。这一现象首次由法国医生 Etienne Lombard 在 1911 年发现并加以描述，此后这种现象被称为 Lombard 效应 (Castellanos et al. 1996; 高慧等 1999)。因此噪声听觉反馈实验一般也被称作 Lombard 实验 (Zhao & Jurafsky 2009)。值得注意的是，同样是研究噪声对言语的影响，噪声听觉反馈研究和一般意义上的噪声中言语传播研究的关注点不同，前者主要关注噪声对说话人本身言语的影响，后者主要关注言语传播过程中噪声对听话人理解效果的影响。

根据以往的研究，与正常条件相比，在 Lombard 条件下言语发声的主要声学变化特征是 (Junqua 1996; Skowronski and Harris 2006)：(1) 基频提高 (increase in fundamental frequency) (2) 能量从低频段向高频段转移 (shift in energy from low frequency bands to middle or high bands) (3) 元音拉长 (increase in vowel duration) (4) 共振峰发生变化 (shift in formant center frequencies for F1 (mainly) and F2) 等。在 Burnett et al. (1997) 的听觉反馈实验中，实验者通过改变反馈音的基频，揭示在人脑神经系统中，基频和音强可能是被独立控制的，它们具有不同的神经机制。那么在噪声反馈条件中，音强的改变到底会带来哪些影响？人脑神经系统对语速的控制与对音高和音强的控制是否运用不同的神经心理机制？这些问题尚需进一步研究。

通过对以往研究的考察可以发现，首先，从研究涉及的语种来看，Lombard 实验涉及西方语言比较多，如英语、法语、西班牙语等，对日语也有很多研究，但对汉语的研究在国际上有影响的成果相对较少。其次，在研究方法上，探讨如何对噪声中语音识别进行计算建模，从而提高识别效果的研究比较多，如 Hansen (1996), Cerisara et al. (2004), Skowronski and Harris (2006)。但是从语言理论视角探讨听觉反馈的语言能力的研究则比较少见。此

---

<sup>②</sup> 在语音学中，发声 (phonation) 专指喉部发声的活动，发音 (articulation) 指在喉以上声道内改变气流以产生各种语音时发声的生理运动 (克里斯特尔 2000: 282, 26)。本文的“言语发声”是指与言语理解 (speech comprehension) 相对的言语产生 (speech production)。

外，以往Lombard实验中的语料尽管也涉及孤立词 (isolated words) 和连续语音 (continuous speech) (Bond 1989, Junqua 1993)，然而，过去的实验要么以单发/a/音为刺激材料，要么以单词或句子作为刺激材料，缺乏通过同一被试，在相同噪声环境下，采用不同刺激材料的汉日对比研究。

汉语属声调语言，其声调的变化主要在于音高 (pitch) 变化。非声调语言的音高变化主要担负着语气功能。汉语音高变化不仅担负语气功能而且担负辨义功能。赵元任曾从理论上指出，汉语音高可以区别词义这一特性使得汉语在噪声传播中是一个优点 (赵元任 2006)。目前这一观点尚未得到实验证实。日语属于音高重音型语言 (pitch-accent language)。它也有丰富的音高变化并且音高变化也可以区别词义。比如“箸 (はし hashi) (HL)”、“橋 (はし hashi) (LH)”等，其音高从高到低 (H-L) 或从低到高 (L-H) 的变化也意味着词义不同。可见，选择汉日两种音高变化比较丰富的语言做比较研究，更加有利于在微观视角下考察语言能力的普遍性和差异性。

总之，本文关注的主要问题是，操汉语者和操日语者在同样噪声 (Lombard) 条件下，阅读不同性质的语料时，其言语速度、音高、音强会发生怎样的变化？这些言语声学特征的变化反映了怎样语音听觉反馈系统的神经心理机制？语义和语境因素对语速、音高及音强会带来怎样的影响？汉日听觉反馈声学参数的会出现怎样的异同及这些异同是否反映两种语言能力的普遍性和差异性。听觉反馈研究对语言习得可以带来哪些启示。

### 3 噪声听觉反馈实验 (Lombard Experiment)

对正常话语而言，我们只知道听觉反馈系统在流畅话语中具有重要作用，但很难确切知道听觉反馈系统是怎样起作用的、具体的工作原理如何。就像一台机器，当它运转良好的时候，我们很难判断哪个部件执行怎样特定的功能。而当机器的某个部件被改变或连接路径被改变时，我们通常可以通过损坏时的运行表现来推测某结构的具体功能。鉴于此，目前噪声听觉反馈研究大多采用下列范式：实验中让被试在听自己话语的同时增加噪声或者让被试在几乎听不到自己声音的噪声环境中说话，藉此考察言语声学参数变化特征以及语言能力的本质特征。

#### 3.1 实验假设

在Lombard实验中，过去的实验要么以单发/a/音为刺激材料，要么单独以单词或句子作为刺激材料，缺乏运用同一被试，在同一噪声环境下，采用各种刺激材料的汉日对比研究。本次实验我们从语言学的角度筛选出三种不同性质的语料。分别是：(a) 持续单发/a/音 (无语义也无语境)；(b) 汉语单词和日语单词 (有语义无语境)；(c) 汉语或日语课文 (有语义有语境)。

汉语声调的变化主要属于音高 (pitch) 变化。赵元任先生认为汉语音高可以区别词义这一特性使得汉语特别便于在嘈杂的环境中传递。非声调语言的音高变化担负着语气功能。汉语中音高变化不仅担负语气功能更主要担负辨义功能。在噪声环境中, 语音的其它特性比如音强、音长等容易受到影响, 而音高则相对稳定。因此, 汉语音高可以区别词义这一特性使得汉语在噪声传播中是一个优点 (赵元任 2006)。

基于此, 我们的实验假设是, 操汉语者和操日语者在同样噪声条件下, 在阅读上述三种不同性质的语料时, 其言语速度、音高、音强会发生不同的变化。汉语和日语两种不同母语背景的被试在相同Lombard条件下会出现不同的言语发声特点。

### 3.2 实验材料及噪声选择

本实验筛选不同性质的语料作为刺激材料, 目的是为了对比研究。刺激材料如下:

汉语被试	日语被试	刺激材料性质
a. 持续单发/a/音	持续单发/a/音	(-语义 -语境)
b. 汉语单词朗读	日语单词朗读	(+语义 -语境)
c. 日语单词朗读	汉语单词朗读	(+语义 -语境)
d. 汉语课文朗读	日语课文朗读	(+语义 +语境)

基于语言学视角这些材料可分为三个类别: 材料一 (a) 为/a/音发声 (phonation), 它既无语义 (semantic) 也无语境 (context)<sup>③</sup>; 材料二 (b, c) 为单词, 它有语义无语境; 材料三 (d) 为课文, 它既有语义又有语境。全部实验语料详见附录。

关于噪声选择, 以往Lombard实验选择的噪声各不相同, 有的选白噪声 (white noise), 有的选粉红噪声 (pink noise)、有的选择自行录制合成的多人谈话言语噪声 (Babble noise)。而且研究业已表明, 白噪声、粉红噪声、言语噪声等不同噪声对言语识别的影响是不一样的 (卢华平等 1996)。本实验选择了70分贝的粉红噪声作为反馈噪声。实验中噪声条件和非噪声条件随机呈现。

共有18例母语为汉语的被试参与实验 (7例为女性, 年龄23-42周岁, 平均28.44周岁 (±5.40))。他们均为北海道大学的中国留学生。有18例母语为日语

<sup>③</sup> 此处/a/音为持续时间为6秒的发声的视觉提示, 它与单词发音以及阅读句子比较而言, 是无语义内容的, 本实验不考虑单独的元音和辅音的发音问题。国际上在相关的实验中, 一些学者在利用/a/音作为发声任务时均认为它是无语义的材料, 详见Kawahara et al. (1996), Burnett et al. (1997), Lincoln et al. (2006) 等论文。另, 这里的语境指上下文 (context)。

的被试参与实验(7例为女性, 年龄20–30周岁, 平均22.83周岁(±3.37)), 均为北海道大学的日本学生。所有被试均为右利手, 无说话、阅读和听觉障碍, 也未受过专业系统的声乐训练。被试均自愿同意参加实验, 每个被试实验大约持续40分钟, 实验后得到一定的报酬或奖励。

实验过程中, 被试单独在一间具有隔音效果的录音室里; 实验中采用坐姿, 即坐在椅子上看着面前的电脑屏幕, 并按要求进行自然阅读。

实验指导语有两条。日语被试用日语提示, 汉语被试用汉语提示。

汉语为: a. 请您按照屏幕显示的内容清楚地将其读出来。b. 读得快慢无所谓, 像平常一样轻松地读即可。

日语为: a. 画面の文章をはつきりと読んでください。b. 落ち着いて普段通りの速さで読んでください。話すスピードは速くても遅くても構いません。

### 3.3 实验任务

被试的实验任务有如下四项, 实验中每项任务的正常条件和噪声条件都重复五次。

(1) 被试看到屏幕上出现/a/时, 就开始持续自然地发/a/音, 时间6秒。

(2) 日语单词朗读。日语单词分别精选了音高变化从高(H)到低(L)变化和从低(L)到高(H)的变化可以导致意义不同的词: “海(umiうみ)、膿(umiうみ)、熊(くまkuma)、隈(くまkuma)、鮭(さけsake)、酒(さけsake)、隅(すみsumi)、炭(すみsumi)、鑿(のみnomi)、蚤(のみnomi)、箸(はしhashi)、橋(はしhashi)、蛎(かきkaki)、垣(かきkaki)、朝(あさasa)、麻(あさasa)、赤(あかaka)、垢(あかaka)、白(しろshiro)、城(しろshiro)”。

(3) 汉语单词朗读。汉语单词挑选了普通话中四个不同的声调可以区别意义的词: “搭(dā)、答(dá)、打(dǎ)、大(dà)、摸(mō)、魔(mó)、抹(mǒ)、墨(mò)、裸(kē)、壳(ké)、渴(ké)、客(kè)、啊(ā)、啊(á)、啊(à)、啊(à)、衣(yī)、姨(yí)、椅(yǐ)、艺(yì)、乌(wū)、吴(wú)、五(wǔ)、雾(wù)、迂(yū)、鱼(yú)、雨(yǔ)、玉(yù)、妈(mā)、麻(má)、马(mǎ)、骂(mà)、咪(mī)、谜(mí)、米(mǐ)、蜜(mì)、督(dū)、犊(dú)、堵(dǔ)、杜(dù)、八(bā)、拔(bá)、靶(bǎ)、爸(bà)”。

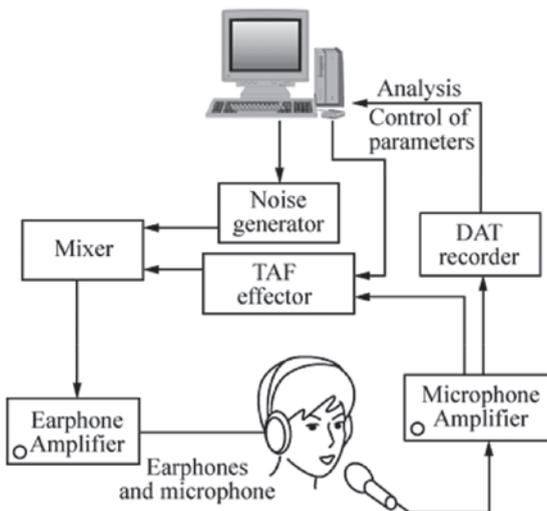
(4) 汉语/日语课文朗读。阅读材料选自《哈利·波特》(Harry Potter) 汉语版或日语版。中国人读汉语, 日本人读日语。阅读材料在电脑屏幕上显示时间为15秒, 15秒后立即消失, 被试随之立即停止阅读, 不管读到句子中间或结尾都立刻停止。实验前被试被告知不必根据记忆把句子读得完整, 这样便于统计单位时间内的阅读速度。阅读材料见附录一。

其中第2项和第3项内容的次序针对汉语被试和日语被试有所不同。针对汉语被试, 第2项为读日语单词(非母语), 第3项读汉语单词(母语); 针对日

语被试，第2项为读汉语单词（非母语），第3项为读日语单词（母语）。实验中有的日语被试学过一段时间汉语，所以对汉语单词的阅读并不困难，有的日语被试以前没接触过汉语，我们就在实验正式开始前进行训练，直到被试报告他自己会读为止。训练中我们也发现，汉语声调尤其是第三声对日语被试而言有一定的难度。汉语被试由于在日本留学，之前或留学期间或多或少都懂一些日语。针对被试不同日语水平，我们对其进行训练，直到被试自己报告可以参加实验为止。有的汉语被试的日语相当好，读之前只要提醒他们注意音高的变化即可。整个发音（phonation）及阅读时间（包括停顿时间）大约22分钟。

### 3.4 实验设备及框架

实验设备及框架如图(1)所示。实验包括三个组成部分：录音装置、听觉反馈改变装置和噪声掩蔽装置。其中录音装置包括心形指向性话筒、放大器、录音设备等；听觉反馈改变装置是指电脑控制的用来改变语声的设备，包括电脑、声音放大器、音乐设备数字接口(MIDI)、A/D转换器(Yamaha Uw 500)、效果器(RFX-2200)等；噪声掩蔽装置是指采用噪声来掩蔽骨传导效应。实验中，发音人口部与话筒之间的距离为6厘米。录音室的本底噪声 $\leq 25$ 分贝(dB SPL)。录音软件是adobe audition 1.5。



图(1) 实验流程框架（基于 Jones and Munhall 2002）

## 3.5 数据分析

为考察被试的言语表现，我们计算了言语的音高变化 (pitch change)、言语速度 (speech speed) 以及平均音强的变化 (power mean) 三个指标。需强调的是，本实验的数据来自多个发音人 (18例汉语18例日语)，因此所得结果为统计学意义的结果。这也是朱晓农强调和推荐的做实验语音学的方法 (朱晓农2005: 340)。朱晓农曾指出“如果随意用一些语图，那么你想说明什么都是可能的，因为一个语言信号有无穷多种随机变异的物理形式——最近二三十年来语音学的最大成就也许就是从实验上证明了这一经验常识”。(朱晓农 2006: 4, 51) ④

### 3.5.1 平均音高

分析时首先通过软件“音声録聞見for windows” (日语版) 把录下来的时域波形 (original) 的基频抽取出来，然后通过自己编制的Matlab程序计算2毫秒之内的基频变化率。

具体操作步骤是先利用“音声録聞見for windows”软件打开录下来的言语声波后 (图(2))，接下来首先设定一定的参数，然后执行 (execute) ⑤。最后点击“文件”中的“另存为”就可以把基频的具体数据以EXCEL形式保存下来 (Excel文档第4列即D列为基频的数值，请参见下文图3)。这样再通过MATLAB计算就能比较容易得到2毫秒之内的基频变化值 (图4)。最后取平均值用来比较。Matlab计算程序请参看附录二。

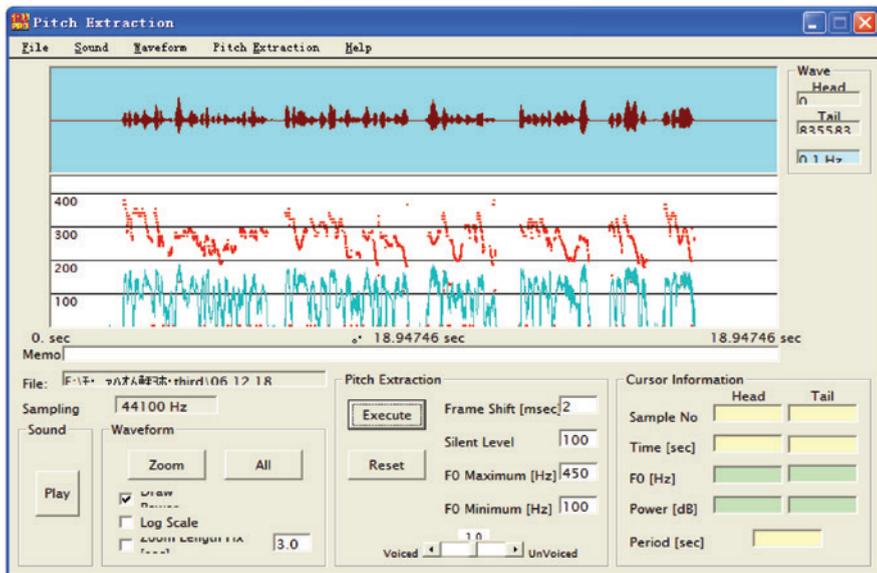
男性的言语频率一般比女性低 ⑥，并且经过试分析发现，汉语被试和日语被试的频率值一般也不一样。在考虑不同性别和不同语言等因素后，在采用“音声録聞見for windows” (日语版) 软件抽取基频时的参考值一般设定如下：

操汉语者 (性别：男)	操汉语者 (性别：女)
Frame shift 2	Frame shift 2
Silent level 100	Silent level 100
F0 Max 400	F0 Max 450
F0 min 60	F0 min 100

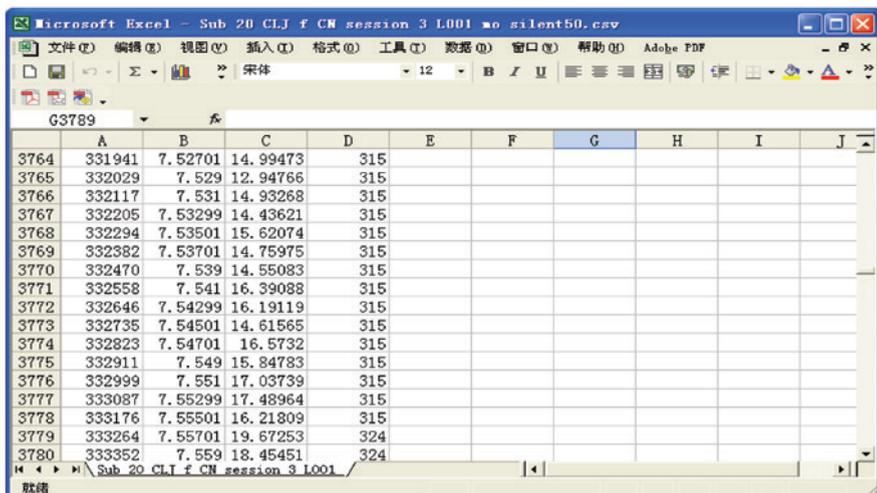
④ 与以往相比，近年来语音学国际期刊所发表论文中有两个较为显著的变化：一是从单个发音人到多个发音人的变化；二是各种统计方法的广泛使用，可参看Jones & Munhall (2002)、Skowronski & Harris (2006)、郑秋豫 (2008)、Zhao & Jurafsky (2009) 等论文。

⑤ 这时候的基频中可能还包含噪声，为了去噪，我们再重新调整基频的最大值和最小值，然后重新执行，直到得到最佳效果。

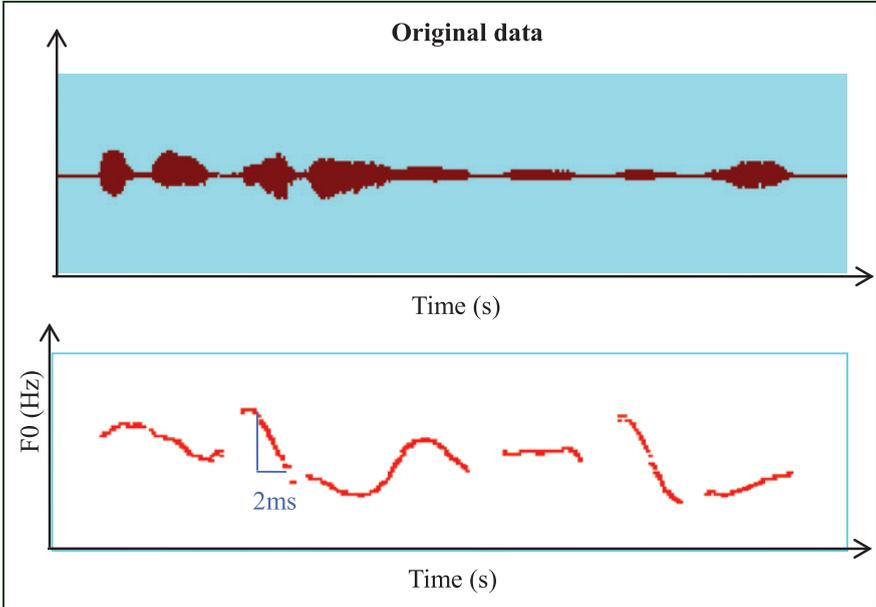
⑥ 男性的F0范围一般在60–240 Hz；女性的范围在180–400 Hz (请参看Cruttenden 2002: 3, 李爱军 2005)。



图(2) “音声録聞見for windows”软件频率抽取演示



图(3) 基频和音强变化的数据说明



图(4) 计算2毫秒之内的基频变化率演示

操日语者 ( 性别 : 男 )	操日语者 ( 性别 : 女 )
Frame shift 2	Frame shift 2
Silent level 100	Silent level 100
F0 Max 300	F0 Max 400
F0 min 60	F0 min 100

### 3.5.2 平均音强 (power mean)

首先利用软件“音声録聞見for windows”(日语版)把录下来的时域波形的音强抽取出来，然后用Matlab去处噪声后计算每个参数的音强平均值。

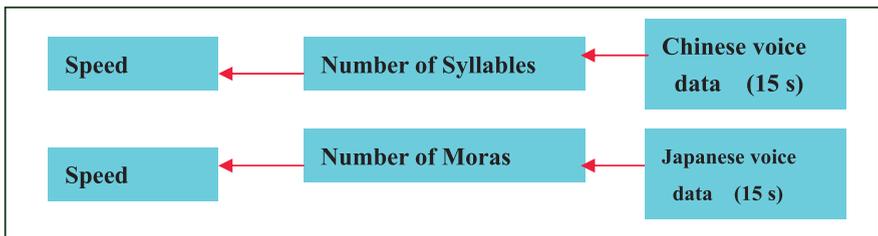
在具体操作中我们利用“音声録聞見for windows”软件打开录下来的言语声波，接下来先设定一定的参数，然后执行 (execute)，请参看上图(2)。完成后点击“文件”中的“另存为”就可以把基频以及音强变化的数据以EXCEL形式保存下来 ( 上图2 )。Excel图 ( 图3 ) 中第3列 ( C列 ) 为音强 (power) 的具体值，第4列 ( D列 ) 为基频 (F0) 的具体值，第1列 ( A列 ) 为采样点 (sample no.)，第2列 ( B列 ) 为时间点 (time)。这样再通过Matlab计算就能比较容易得到2毫秒内基频变化值以及音强平均值。

由于男性的言语频率一般比女性低，并且汉语被试和日语被试的基频值也不一样，本实验中采用“音声録聞見for windows”软件时参考值设定见3.6.1。

需要说明的是，与上面DAF和FAF实验分析相比，这次的“silent level”值一般都设为50，因为有的被试发/a/音时，如果“silent level”值设为100，基频根本看不到。根据实际分析的需要，有的个别被试（比如汉语被试No.20）“silent level”值需设得更小（为20）才能保证分析效果。

### 3.5.3 言语速度

计算15秒内被试阅读的音节（汉语）或者摩拉（日语）数量（图5）。由于汉语是以音节 (syllable) 为单位的语言，基本上一个音节代表一个汉字，因此，本文采用单位时间内被试阅读音节的数量来衡量汉语被试的阅读速度；日语是以摩拉 (mora) 为单位的语言，因此本文采用单位时间内被试阅读的摩拉数量来衡量日语被试的阅读速度<sup>⑦</sup>。



图(5) 言语速度计算

### 3.5.4 比率换算

在计算中以2毫秒为间隔计算出录音的基频( $F_0$ )曲线（采样频率44100 Hz），把这些以2毫秒为单位的计算位置定义为特征取样点，并手工去除曲线上因自动算法提取错误而造成的奇异点，然后通过Matlab计算每一特征取样点与前一特征取样点的基频差值。设当前特征取样点基频值为 $f_i$ ，前一特征取样点的基频值为 $f_{i-1}$ ，则对当前特征取样点得出 $\Delta f_i = f_i - f_{i-1}$ ，最后，对所有特征取样点的 $\Delta f_i$ 的绝对值取平均值，定义为基频的平均绝对差值。针对言语速度，计算每15秒内被

<sup>⑦</sup> 关于摩拉 (mora) 的含义及更深刻理解，请参看Gussenhoven and Jacobs (2001: 160–164) 等。

试阅读音节（汉语）或者摩拉（日语）数量；针对平均音强，先利用软件“音声録聞見for windows”以2毫秒为间隔把录音的音强抽取出来，然后用Matlab计算每个参数的音强平均值（与音高平均值计算类似）。

为便于比较汉语和日语之间异同，不管是2毫秒间隔的平均音高变化、15秒内所读音节或摩拉数，还是平均音强数据，均被换算成相应的比率。换算公式如图(6)：

变化率 (ratio) = $\frac{\text{正常条件下的量}}{\text{改变条件下的量}}$	正常条件下的量	Ratio (e.g.)
	改变条件下的量	$150 / 100 = 1.5$ $120 / 100 = 1.2$

图(6) 变化率转换计算（这里的量包括基频的平均绝对差值、音强的绝对差值以及平均音强/摩拉数）

由于变化率为正常条件值除以改变条件值，因此变化率越大，影响（包括音高和语速和音强）也就越强。以操汉语者言语速度为例，假定正常条件所读音节数为150，改变条件为100，则比率等于1.5；相比另一条件，正常条件音节数为120，改变条件为100，则比率等于1.2；1.5的比率大于1.2的比率，说明前一种条件下所受的影响要更加强烈。

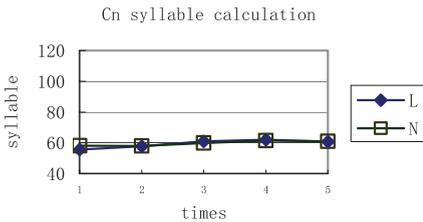
### 3.6 实验结果

需说明的是，实验中操汉语者的被试实际上是20例。但是编号为“sub 15”的被试的计算结果与平均值相差很大，复查其背景显示<sup>⑥</sup>，该被试为心理学专业研究生，并且也是做心理语言学研究的。本文认为该被试的专业背景可能对实验结果构成影响，所以该被试最后统计时被排除在外。此外，编号为“sub 17”的被试结果由于出现和“sub 15”类似的倾向，最后统计时也被排除在外。下面是18例汉语被试和18例日语被试的Lombard实验结果（所有结果均经*t*检验。 $p < 0.01$ 。见图(7-15)和表(1-2)。

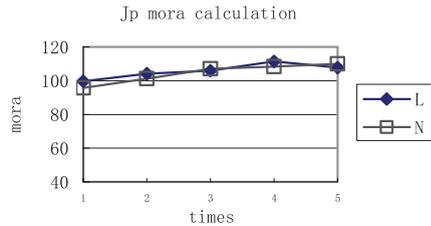
⑥ 本文研究中的每一被试在实验后均填写一张问卷调查表，表中包括可能与研究有关的各种背景信息。

### 3.6.1 组内比较 ( 正常条件与噪声条件 )

(1) 正常与噪声条件对比，日语和汉语被试各自在言语速度上都没有显著差异。

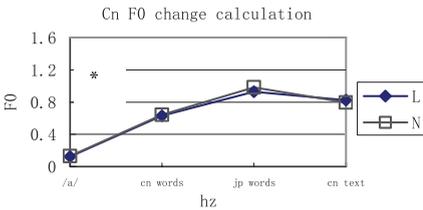


图(7) 汉语正常与Lombard条件语速比较 (n = 18)

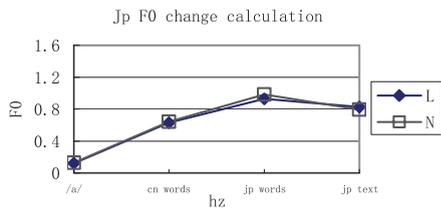


图(8) 日语正常与Lombard条件语速比较 (n = 18)

(2) 正常与噪声条件对比，日语和汉语被试在音高变化 (pitch change) 上都没有显著差异 (其中汉语被试在/a/发声的任务除外， $p = 0.0301$ )。

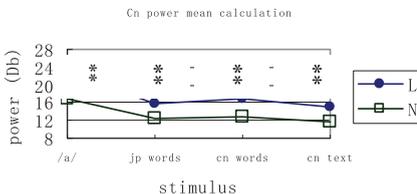


图(9) 汉语正常与Lombard条件音高比较 (n = 18)

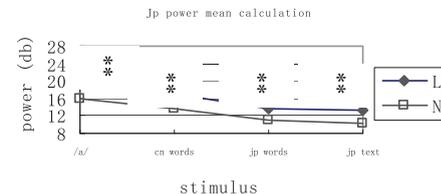


图(10) 日语正常与Lombard条件音高比较 (n = 18)

(3) 正常与噪声条件对比，日语和汉语被试在音强 (power mean) 上都有显著差异。



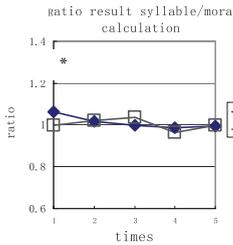
图(11) 汉语正常与Lombard条件音强比较 (n = 18)



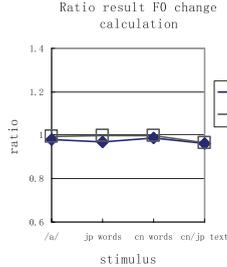
图(12) 日语正常与Lombard条件音强比较 (n = 18)

3.6.2 组间比较 (汉语和日语被试语速、音高及音强变化率之间的比较)

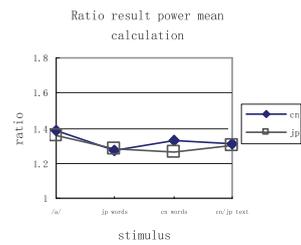
日语和汉语被试在言语速度变化率上均无显著差异 (其中第一次条件除外,  $p = 0.0188$ )。在音高和音强变化率上也无显著差异。



图(13) 言语速度变化率比较



图(14) 音高变化率比较



图(15) 音强变化率比较  
(n = 18 each)

表(1) 组内不同条件之间比较 (/a/ 音、单词、课文) (\*\*表示显著意义, no表示无显著意义)

cn power mean		/a/ vs jp words	/a/ vs cn words	/a/ vs cn texts	jp words vs cn texts	cn words vs cn texts
	N	**	**	**	no	*
	L	**	**	**	no	**
jp power mean		/a/ vs jp words	/a/ vs cn words	/a/ vs jp texts	cn words vs jp texts	jp words vs jp texts
	N	**	**	**	**	no
	L	**	**	**	**	no
cn F0 change		/a/ vs jp words	/a/ vs cn words	/a/ vs cn texts	jp words vs cn texts	cn words vs cn texts
	N	**	**	**	**	**
	L	**	**	**	**	**
jp F0 change		/a/ vs jp words	/a/ vs cn words	/a/ vs jp texts	cn words vs jp texts	jp words vs jp texts
	N	**	**	**	**	**
	L	**	**	**	**	no

表(2) 组内不同条件之间比较 (/a/ 音、单词、课文) (\*\*表示显著意义, no表示无显著意义)

N power mean		/a/ vs jp words	/a/ vs cn words	/a/ vs cn/jp texts	jp words vs cn/jp texts	cn words vs cn/jp texts
	cn	**	**	**	**	no
jp	**	**	**	**	no	**
L power mean		/a/ vs jp words	/a/ vs cn words	/a/ vs cn/jp texts	jp words vs cn/jp texts	cn words vs cn/jp texts
	cn	**	**	**	no	**
jp	**	**	**	**	no	**
N F0 change		/a/ vs jp words	/a/ vs cn words	/a/ vs cn/jp texts	jp words vs cn/jp texts	cn words vs cn/jp texts
	cn	**	**	**	**	**
jp	**	**	**	**	**	**
L F0 change		/a/ vs jp words	/a/ vs cn words	/a/ vs cn/jp texts	jp words vs cn/jp texts	cn words vs cn/jp texts
	cn	**	**	**	**	**
jp	**	**	**	**	no	**

### 3.6.3 组内不同条件间比较 (/a/ 音、单词、课文)

根据表(1-2)得到的结果如下：(1) 被试在单发/a/时以及被试在读单词或读课文时，无论在正常条件还是在噪声条件下，言语的音高和音强两个指标变化是不同的。这显示如果在言语发声中增加了语义、语境因素，它们对音强(power)以及基频(F0)都会产生显著影响。(2) 在音高变化(pitch change)上，除日语被试在读日语单词与日语课文之间无显著意义之外(仅在噪声条件下,  $p = 0.25$ )，无论在有无噪声的条件下，两种语言被试在读单词和读句子时，在音高变化上都有显著差别。(3) 在平均音强(power mean)上：(a) 正常条件下，对汉语被试而言，汉语单词与汉语课文之间有显著差异，日语单词与汉语课文之间无显著差异；对日语被试而言，汉语单词与日语课文有显著差异，日语单词与日语课文无显著差异。(b) 在Lombard条件下，趋势与正常条件的一样。

综上，考虑本文Lombard实验涉及的因素较多这里先把Lombard的实验结果作一小结。

(一) **组内比较** : (1) 正常与噪声条件对比, 日语和汉语两组被试各自在言语速度 (voice speed) 上都没有显著差异。(2) 正常与噪声条件对比, 日语和汉语两组被试各自在音高变化 (Pitch change) 上都没有显著差异。(3) 正常与噪声条件对比, 日语和汉语两组被试各自在音强变化 (Power mean) 上都有显著差异。

(二) **组间比较** (正常条件与改变条件变化率之间的比较) : (1) 日语和汉语被试在言语速度 (speech speed) 变化率上没有显著差异。(2) 日语和汉语被试在音高 (pitch change) 变化率上没有显著差异。(3) 日语和汉语被试在音强 (power mean) 变化率上没有显著差异。

(三) **组内不同条件间比较** (/a/音、单词、课文) : (1) 被试在单发/a/音时以及被试在读单词或课文时, 无论在正常条件还是在噪声条件下, 言语的音高和音强两个指标变化是不同的。这显示在言语发声中如果增加了语义、语境因素, 它们对音强 (power) 和基频 (F0) 会产生显著影响。(2) 在音高变化 (pitch change) 上, 除日语被试在日语单词与日语课文之间无显著意义之外, 无论在有无噪声的条件下, 单词和句子之间都有显著差别。(3) 在平均音强 (power mean) 的差异上, 无论是在正常条件还是在Lombard条件下, 对汉语被试而言, 汉语单词与汉语课文之间有显著差异, 日语单词与汉语课文之间无显著差异; 对日语被试而言, 汉语单词与日语课文有显著差异, 日语单词与日语课文无显著差异。

## 4 讨论

### 4.1 阅读速度、音高和音强变化及反映的神经心理机制

关于Lombard实验, 已有的研究显示, 对正常听力者和耳聋者而言, 当言语信号受到噪声干扰时, 言语识别率都将受到不同程度的影响 (Studebaker et al. 1994)。那么在同样的噪声条件下, 当日语被试和汉语被试阅读自然语句时, 他们的阅读速度、音高变化、平均音强等会有哪些异同呢? 这些异同又反映怎样的语言现象呢? 接下来将针对这些问题结合已有研究成果及本文实验结果进行讨论。

本实验对被试每项任务 (/a/音、单词、课文) 的正常条件和噪声条件都重复五次。计算结果时采用五次测试的平均值。在被试读课文的言语速度上, 汉语被试五次的平均结果为: 正常条件下15秒内所读音节 (syllable) 数为59.63, 噪声条件下15秒内所读音节为59.36。可见, 汉语被试在正常条件下比在Lombard条件下读得要稍快一些, 说明噪声条件对言语速度还是有一定的影响。但是 $t$ 检验结果显示, 这种差异并不具有统计学上的显著意义, 说明噪声条件对言语速度的影响并不显著。日语被试五次阅读的平均结果为: 正常条件下15秒内所读摩拉 (mora) 数为104.49, 噪声条件下15秒内所读摩拉数为

105.78。可以看出噪声条件下阅读的摩拉数比正常条件还略高些，但是 $t$ 检验结果显示，这种差异同样并不具有统计学上的显著意义，说明噪声条件对日语被试阅读速度基本不构成显著影响。总之，综合两种语言被试的结果可以看出，噪声条件对言语速度的影响并不具有显著意义。

卢华平等(1996)的实验结果表明，对正常听力者和耳聋者而言，当言语信号受到噪声干扰时，言语识别率会受到影响。本文实验结果与此结论并不一致。可能的原因是：第一，实验中噪声种类不同会导致结果的不同，本实验采用的是粉红噪声，而卢文采用的是言语噪声。第二，卢文的实验是根据词汇来判断识别率，本文是考察单位时间内自然语句的阅读速度。可以说，实验任务的不同可能导致结果的不同。从语言认知的角度来看，人脑神经系统对语言理解具有一定的容错性能力(Saberi and Perrott 1999)，与词汇识别任务相比，在读自然语句的时候，被试可以利用上下文等语境因素来提高言语识别率。也就是说，在有语境的条件下，人脑听觉神经系统对语言理解的容错能力得到提高。因此，噪声对言语速度的影响需要做进一步分析：有无语境内容也可能导致不同的表现和结果。

Burnett et al. (1997) 在听觉反馈实验中让被试(共67位)持续发/a/音，在发音过程中被试反馈音的基频被提高或降低，实验中被试被告知尽量忽略反馈音信号并保持自己的发声在自然状态。实验的目的是观察被试的言语声学特征的变化。结果发现，当反馈音的基频降低时，96%的被试的言语基频会提高，当反馈音的基频提高时，78%的被试的言语基频会降低，即被试的言语基频会出现补偿性反应(compensating response)；但是，在两种条件下，被试言语的音强均无相关变化。基于此，研究者推测，神经系统对音高和音强的控制是相互独立的，它们可能具有不同的神经机制。但该实验并没有涉及语速与音高以及音强是否相互独立问题。

本实验中，**在课文阅读速度上**，汉日两种语言被试的结果均显示噪声条件对阅读速度的影响**不具有显著意义**。

**在音高变化上**，本文测试结果表明，无论对何种任务条件而言(/a/音、单词、课文)，日语和汉语被试在音高变化上均无显著差异(图9-10)。说明正常条件下和噪声条件下的区别都**不具有显著意义**。此结果与Wassink et al. (2007) 所归纳的结论一致。

**在平均音强的变化上**，无论是对哪种任务而言(/a/音、单词、课文)，正常条件与噪声条件下的差别都具有极显著意义(图11-12)。说明Lombard 条件对音强的影响十分强烈。

对比Lombard条件对速度、音高和音强的影响结果可以发现，它对语速和音高的影响不具有显著差异，但对音强的影响却具有显著差异。这个结果不仅再次证实神经系统对**音高和音强**的控制是相互独立的，而且，该结果进一步证实神经系统对**语速和音强**的控制也可能是相互独立的，人脑神经系统对语速和音强的控制可能运用不同的认知心理机制。

语音的音高、语速和音强等在神经系统中被独立控制的属性在语言传输中具有一定的优点。根据Moore (2004: 329) 的研究，像噪声或频谱的改变等每种

干扰条件一般只会影响到语音的某一特性，但是声波的其它特性仍然可以传输信息，这在实践中是个很大的优点。假如言语感知依赖单一因素，那么言语交际在大多数现实环境中将会变得相当困难。

此外，本实验结果还验证了赵元任提出的汉语在噪声传播中具有一定优势的观点。汉语属声调语言，它在一定程度上靠音高变化来区别意义。对非声调语言而言，音高变化主要担负语气功能，但对汉语而言，音高具有双重作用，它不仅担负语气功能而且担负辨义功能。赵元任曾从理论上指出，汉语音高可以区别词义这一特性使得汉语在噪声传播中是一个优点。但是该观点长久以来一直没有实验证实。本文的实验结果揭示，噪声条件仅对语音的音强影响比较显著，而对音高的影响并不显著。这也就是说，在噪声中语言的音高相对稳定。因此，汉语音高的感知在噪声条件下相对稳定。所以可以说，“汉语音高可以区别词义这一特性使得汉语在噪声传播中是一个优点”。基于以往的研究以及本文的分析结果表明，人脑神经系统对音高、音强和语速的控制具有相互独立性，因此，汉语在噪声传播中的稳定性特点具有一定的认知心理基础。

## 4.2 语义和语境因素对语速、音高及音强的影响

从以往的文献来看，过去听觉反馈的实验材料主要是运用持续的 /a/ 音 (sustained /a/ phonation) 或者单词 (Kawahara et al. 1996, Burnett et al. 1997)。这样的实验材料既不含语义也不含语境。通过这类性质的材料来推测言语发声的听觉反馈机制存在一定缺陷。因为实际的言语发声都是既含语义也含语境的自然语句。根据 Lincoln et al. (2006) 的考察，随着研究进展，在过去 10 年的时间里，已经出现了不少以句子为阅读材料的实验研究。如 Natke et al. (2001) 在实验中已经考虑到采用自发言语 (spontaneous speech) 作为实验语料。他们在实验中的具体做法是选取一些主题 (topic)，然后让测试者围绕该主题自发讲话。这样的研究由于考虑到自然言语中的语义和语境因素对语速、音高及音强的影响，因此实验结果更具说服力。

根据王士元等著作的报道，有语义内容和无语义内容的发音其神经系统活动区域是不一样的。如在一项实验中，被试在念 [p] 的时候，比如读以 [p] 开头的单词 (plash、pledge 等)，其嘴唇的动作和吹蜡烛或吹一张纸的动作很相像，但是这两种相同的动作激活的脑波是不同的。吹纸的时候左脑和右脑的波形没有什么区别，但在发 [p] 音时，左脑听觉诱发电位 (AEP, Auditory Evoked Potential) 非常大，右脑没有什么变化 (王士元、彭刚 2006: 35)。导致该差异的原因在于前一个动作属有语义内容的发音，而后一个动作属无语义内容的发音。可见，有无语义内容被试发音的神经基础是不同的。无论在正常环境还是噪声环境下，语义及语境因素对音强和音高变化的影响是显著的。这个结果与以往关于元音感知的研究结果一致 (Keith 2005)。

但是，在听觉反馈中，语境因素是否和语义因素一样可以对言语的语速、音高及音强产生影响？有无语境因素发音的认知心理基础是否不同？这都需要进一步的研究来证明。

本实验中，通过组内不同性质参数（/a/音、单词、课文）的比较结果显示（表(1)、(2)）：被试在单发/a/音以及在读单词（+语义）或课文（+语境）时，无论在正常还是在噪声条件下，言语的音高和音强两个指标变化均具显著性<sup>⑨</sup>。这显示在言语发声中如果增加了语义、语境因素，它们对音强（power）和基频（ $F_0$ ）的变化会产生很大影响。

因此，本文Lombard实验结果进一步证实，无论在正常还是噪声环境下，当增加语义、语境因素之后，说话人的发音在音高变化、平均音强上均具显著差异。也可以进一步推断，有无语义和语境因素，被试言语发声（speech production）的神经心理机制是不同的。

### 4.3 汉日听觉反馈声学参数的异同及两种语言能力的普遍性和差异性

世界上现有的语言约有三千多种，它们之间存在着语音、语义或语法等各种差异。然而，人们通过大量的研究发现，在这些差异的背后有很多相同或相近的共性存在于所有语言当中。Hockett (1966) 曾经指出：“如果以动物的交际系统为参照的话，那么，所有的人类语言都有如下的共同特征：(1) **发音-听觉渠道（都以发音-听觉为渠道）**；(2) 散射传播和有向接收；(3) 快速消失；(4) 互换性（所有的人都同时是语言信号的传播者和接收者）；(5) **反馈性（语言信号的传播者本人接收到信号）**；(6) 专门性；(7) 语义性；(8) 任意性；(9) 离散性；(10) 移置性（displacement）（语言信息可以指在时间、空间上远离交际地点的事物）；(11) 开放性；(12) 继承性；(13) 双重性（既有形式子系统又有意义子系统）；(14) 可仿性（语言信息有虚假的可能）；(15) 反身性（人们可以通过谈论交际而交际）；(16) 可学性（说某种语言的人可以学得另一种语言），等等。”可见，反馈性是世界语言的共性之一。但是，其实在语言学界对语言反馈性的研究不够充分重视。

Hauser, Chomsky and Fitch (2002) 在《科学》(Science) 上的文章中提出语言能力可分为广义和狭义两种。广义语言能力包括感知-运动系统（sensory-motor system）、概念-意图系统（conceptual-intentional system）和递归计算系统（computational mechanisms for recursion）。虽然该文并未明确指出听觉反馈系统是否属于广义语言能力，但从其研究内容上看，由于听觉反馈研究主要

<sup>⑨</sup> 由于本文采用的是单位时间内阅读的音节或摩拉数来衡量言语速度，因此言语速度的比较只在句子阅读任务中进行，而此处比较涉及无语义的/a/的发音（phonation），所以这里仅比较音高和音强。

探讨言语感知-发声 (perception-production) 问题, 它也就很自然地成为“感知-运动”系统研究的一个重要组成部分。因此, 语音听觉反馈能力自然也是构成人类语言能力的一个有机组成部分。其实, 语音听觉反馈能力属于人类语言能力的重要组成部分这一观点从结构主义语言学思想中便可发现端倪。索绪尔特别强调说和听的统一, 其中的一个要有另外的一个才能有它的价值, 如果我们发出一个音, 耳朵听不到, 就不能形成音响印象 (索绪尔2003: 29), 基于此, 杨亦鸣 (2007) 指出, **听觉系统**也进入了**语言能力**的范围。综合上述学者的研究可以发现, 语音听觉反馈能力是构成人类语言能力的一个重要组成部分。接下来我们将根据实验来分析操汉语者和操日语者的语音听觉反馈能力的普遍性和特殊性。

语言的普遍性和特殊性现象受到生成语言学和语言类型学共同的关注。“国际上有影响的语言学派大多把主要的注意力放在对语言共性的提取上, 而把发掘语言之间的各种差异放在了次要的、服从的理论地位上。而在中国, 对语言共性的研究至今还没有形成规模, 成绩更是微乎其微” (程工1999)。王士元 (2006) 也曾指出, “语言的研究最重要的是一种一般规律性的研究”。

在生成语法学派看来, 像移位 (movement) 等规则在不同语言中具有一定的普遍性, 但普遍性其实可以体现在更深的层面, 比如在语言和思维系统接口的心理表征层面体现。彭聃龄 (2004) 也指出, 语言的普遍性和差异性不仅存在于语言学层面, 也存在于语言的信息加工层面和语言的神经机制层面, 坚持普遍性和差异性的统一, 不仅对语言本身的研究很重要, 而且对了解人脑加工语言的特点也有重要意义。

纵观学界对语言普遍性的研究状况, 目前对许多句法规则或其它语言现象 (比如语序) 的普遍性探讨比较多且较为深入, 而对语言信息加工的神经心理机制层面普遍性的研究则不够深入。在语言的神经心理机制层面, 一方面, 体现人类语言能力的语言神经心理机制具有一定的普遍性, 另一方面, 人类语言神经心理机制具体到各个语言时, 又会由于不同语言文化的差异而表现出一定的个性。比如, 根据以往对阅读障碍 (dyslexia) 的研究表明, 不同语言的阅读障碍具有共同的神经认知基础 (Paulesu et al. 2001, Eden and Moats 2002)。Siok et al. (2004) 针对汉语阅读障碍的被试研究又揭示, 汉语阅读障碍儿童发生功能异常的脑区和西方拼音文字阅读障碍者明显不同。这个结果又体现出不同语言文化背景的阅读障碍发生机理的个性。

在语音层面, 尽管各种语言都利用声音来传达意义, 但每种语言在利用声音的时候都有自己的选择。通过对不同语言的考察表明, 语言在选择声音表达意义时既体现普遍性又体现特殊性。比如各种语言的语音中都包含元音, 这是人类语言的共性, 但是这些元音一旦和重音、声调等因素进行不同方式的组合, 又使各种语言的语音出现了丰富多彩的变化。Whalen and Levitt (1995) 通过考察指出, 每种语言都有这样一致的倾向, 就是高元音比如[i]和[u]比低元音比如[a]的基频要高。这个特征是该文考察31种不同类型的语言之后发现的共有特点。经过学者们的进一步考察发现, 这个普遍性同样适用于声调语言包括非洲的声调语言 (Connell 2002)。

Xu Yisheng et al. (2006) 做了一项操汉语和操英语者的类别感知 (categorical perception) 对比实验。实验中被试的任务是判断言语音 ( 第一声 yī, 衣, speech stimuli ) 和非言语性质的和谐音 (harmonic tones, nonspeech stimuli)。实验中的和谐音具有与言语音同样的音高、音强和音长。结果显示, 操汉语者对言语音的感知结果比操英语者好, 即使对非言语音, 操汉语者的感知结果也比操英语者好。但是, 操英语者对非言语音的感知结果比对言语音的感知结果要好。该文指出, 在短期记忆上不同语言的被试具有共同性, 但是在长期记忆上, 由于语言经验的不同, 体现了不同语言的被试对类别的感知具有一定的个性。据此作者得出的结论是, 类别感知是全人类共同的能力, 但这个能力又受到长期记忆的影响。

可见语言能力普遍性和差异性也体现于在神经心理层面。当然, 尽管以往的一些论述也昭示了言语听觉反馈能力的重要性, 但这些论文并未采用实验的手段对其进行验证。

接下来我们基于本文的实验结果具体讨论操汉语者和日语者的听觉反馈能力 ( 即元认知能力 ) 所体现的普遍性和差异性。

根据实验结果 ( 表1-2 ), 在Lombard条件下, 可以发现汉日两种语言被试在听觉反馈能力上具有普遍性。在Lombard实验中, 正常与噪声条件对比, 日语和汉语被试在言语速度、音高变化上均无显著差异。然而日语和汉语被试在音强上都有显著差异。显示噪声对音强的影响较大, 而对言语速度和音高的影响较小。这个特点体现两种语言具有普遍性。结果还显示, 无论在正常环境还是噪声环境下, 语义及语境等因素对音强和音高变化的影响都是显著的。从汉日两种语言间的结果来看这种显著差异也体现出共性。

由表(1)和表(2)可见, 无论在正常环境还是噪声环境下, 操汉语者和操日语者在读单词和读课文时表现出的差异是复杂的<sup>⑩</sup>。但这个结果恰好体现了汉语和日语在音强变化上的差异性。具体而言, 无论在正常条件还是在噪声条件下, 汉语被试的汉语单词 (+声调) 和汉语课文之间有显著差异, 日语单词 (-声调) 和汉语课文之间在音强上无显著差异。同样, 无论在正常条件还是在噪声条件下, 日语被试的日语单词 (-声调) 与日语课文之间没有差异, 汉语单词 (+声调) 和日语课文之间有显著差异。请参看下表(3)。

这种差异背后的原因是什么呢? 根据Moore (2004: 228) 对音乐中音高感知的研究, 一个人早期经验的不同、文化背景的不同都可能影响感知的效果。我们推测, 个人早期的语言经验和文化背景也会对言语的音高和音强等的感知产生影响。汉语和日语尽管表面上均具有丰富的音高变化, 但两种语言的音高变化还是有区别的, 首先, 汉语音高变化在一个音节之内发生 (within syllable), 而日语音高变化则是跨越音节的 (between syllables)。此外, 汉语的音高变化

<sup>⑩</sup> 这种复杂性体现在音高变化和平均音强的差异上, 它们的变化规律和趋势并不一致。此处我们仅讨论音强的变化规律, 因为它与讨论的主题 ( 语言能力普遍性和特殊性 ) 密切相关。本文暂不涉及对音高变化的讨论, 但我们在表(1)和表(2)中也已给出了实验结果, 希望能给未来的研究者提供参考。

表(3) 汉语和日语被试的单词与课文之间音强比较 (符号“+、-”分别代表“有无显著差异”)

汉语被试		日语被试	
汉语单词VS汉语课文	日语单词VS汉语课文	日语单词VS日语课文	汉语单词VS日语课文
+	-	-	+

可以区别词义是普遍现象,而日语音高变化区别词义并不像汉语那样具有绝对意义;这从两国的词典体例上亦可见一斑——几乎所有的汉语词典对每一个单词都标出声调,而日语中除《(新)明解国语辞典》等极个别词典以外,极少有词典对日语单词标出具体的音高变化。

由于受汉语经验及文化背景的影响,操汉语者对音高的感知会更加敏锐,而对音强的感知则不那么敏感,随着时间的推移,母语为汉语的说话人则会形成对噪声具有一定的抗干扰能力。相比之下,由于日语音高变化并不像汉语那样具有绝对意义,该语言文化背景不需对音高做出类似汉语的敏锐感知,因此,母语为日语的说话人在噪声中对音高和音强的感知与操汉语者会有所不同。这可能是导致操汉语和日语者之间差异的主要因素。

#### 4.4 语言习得与语音听觉反馈

Elman (1981) 提出,听觉反馈主要目的是对语言长期校准 (calibration),但是它在特定条件下对言语的短期调节控制作用也是很必要的。Jones and Munhall (2003) 也指出,听觉反馈系统不但在说话时对言语进行在线控制 (on-line control),而且在对言语发声和感知的长期校准中也起重要作用。虽然到目前为止本文主要讨论听觉反馈对言语发声的在线控制功能,但这并不是说听觉反馈对言语的长期校准功能就不重要。本文认为,听觉反馈对言语的在线控制功能和长期校准功能之间存在着密切的联系。一方面反馈的在线控制的结果会影响到对言语的长期校准;另一方面,听觉反馈中形成的对言语的长期校准的结果(即内部表征)又会对言语的在线控制产生重要的影响。它们的共同作用在早期的语言习得过程中起重要作用。譬如在发音学习中,有效的反馈是必不可少的一个重要环节(韩纪庆等2004: 161, Toyomura et al. 2007)。

心理语言学中关于语言习得 (language acquisition) 的研究表明:语言理解能力的获得先于表达能力的获得(靳洪刚1997: 38-39)。儿童在出生初期的几个月内虽然不能讲话,但是已经能够通过倾听大人的讲话来获得关于母语的语音特征(Kuhl et al. 1997)。这说明在非常早期的阶段,儿童的语言感知就开始对语言发声产生了一定程度的影响。聋人群体一般会出现“十聋九哑”现象。那是因为听力的消失,使他们缺乏对语言的听觉反馈能力,历经一段时间后,也就影响到他们的语言发声能力。临床研究表明,聋哑人不会说话,有许多人

并不是发音器官异常,更不是所谓舌头的问题。其根本原因是因为聋人自小以来长期根本听不到真实的言语声,无法学习说话。而在正常条件下,儿童在学习/习得语言的阶段,通过自己的听觉反馈,在神经系统中逐步建立/激活关于该种语言知识的内部表征(internal representation),这个内部表征系统在以后的语言使用中起很重要的作用。

语言的发展不仅包括习得一套复杂的语法规则,而且包括习得可以区别意义的音位系统(Kuhl et al. 1997)。值得注意的是,在儿童语言发展过程中,其父母与孩子交流时使用一种叫“指向婴儿的语言”(IDS, Infant-directed speech)。“指向婴儿的语言”也称为“妈妈语”(motherese)或“父母式语言”(parentese),是指父母在与婴儿交流时使用的一种习惯性语言。它与“指向成人的语言”(ADS, adult-directed speech)相比,句法和语义都比较简单(Snow and Ferguson 1997)。根据学者们的研究,指向婴儿的语言在声学特征上主要表现为基频升高、语速变慢、夸张的韵律等(Fernald and Mazzie 1991, Carroll 1994: 324, Kuhl et al. 1997, Dominey and Dodane 2004)。这样的语言具有两种功能:一是更加有效地把语音类别区别开来;二是这种夸张性的显著区别更加有利于儿童的模仿(Kuhl et al. 1997)。

针对指向婴儿的语言特点和Lombard条件下的语言特点,Wassink et al. (2007) 总结出一个更有意义的结论,指向婴儿的语言和Lombard条件下言语在音高和音强的调节上具有相似性。因此本文针对Lombard条件下言语发声所发现的特点对婴儿语言能力的发展具有一定的启发意义。

根据实验结果,在Lombard条件下,被试的言语在速度、音高、音强等方面发生了一些变化。被试言语发生的变化是自己主动适应环境的一种调整,调整的目的是使自己的话听起来更加清晰、更容易被人理解。可以说这个目的与父母对儿童讲话时所持有的目的相似。人类似乎有一种普遍天生的能力,就是用夸张性的语言来对婴儿说话,而且这种夸张性的语言正好契合儿童听觉反馈能力发展的需要。据此,我们可以推论,不管是指向婴儿的语言还是噪声听觉反馈条件下的语言,那些适应性的调整都具有共同的目的,这种共同的目的可能就是言语发声参数变化具有一致性的动因所在。

## 5 结语

根据本文的实验结果以及对以往有关听觉反馈研究成果的分析和讨论,可以把本文的研究结论归纳如下。

一、噪声条件与正常条件相比,噪声条件对言语速度和音高变化的影响并不显著,但是噪声条件对言语发声中平均音强的影响非常显著。在言语发声过程中,人脑神经系统对言语音高、语速和音强的控制可能具有不同的神经心理基础。

二、无论在正常条件还是噪声条件下,语义和语境因素对音强和音高变化等因素的影响是显著的。有语义内容和无语义内容发音的神经心理基础是不同

的，这种不同可能是导致在增加了语义、语境因素之后，说话人在音高、音强上都具有显著差异的重要原因。

三、语音听觉反馈能力是人类语言能力的一个重要组成部分。听觉反馈能力反映的是人类运用语言的元认知能力。汉日两种语言被试在语言元认知能力上既有普遍性又有差异性。语言元认知能力的差异性体现为汉语和日语被试在Lombard条件下言语的音强变化具有不同特点。

四、语音听觉反馈在正常言语发声中具有重要作用。本文在噪声听觉反馈条件下得到的关于言语发声的参数变化的成果，对儿童的语言习得、成人的外语学习等均将具有一定的借鉴作用。

#### 附录一：课文朗读（举例）

##### 《哈利·波特》阅读材料（汉语）

1. 现在哈利又把眼睛贴近望远镜，重新调整了一下焦距观察金星。他低头看着自己的星象图，准备把金星的位置记上去，但是有东西分散了他的注意力；手里的羽毛笔悬在羊皮纸上空停了下来，他眯起眼睛望着下面幽暗的场地，看到六个身影正走过草地。如果他们没有走动，如果月光没有照亮他们的头顶，那么在漆黑的场地上他们是不容易被发现的。虽然离得这么远，可哈利还是莫名其妙地感到，自己认出了那些人当中最矮胖的那个人的走路姿势，那个人好像在带领着这个队伍。

2. 他们站在一个很长的金碧辉煌的大厅一头，地上是擦得光亮鉴人的深色木地板。孔雀蓝的天花板上镶嵌着闪闪发光的金色符号，不停地活动着、变化着，像是一个巨大的高空布告栏。四面的墙壁都镶着乌黑油亮的深色木板，许多镀金的壁炉嵌在木板里。每过几秒钟，随着噗的一声轻响，就有一个巫师从左边某个壁炉里突然冒出来。而在右边，每个壁炉前都有几个人在排队等着离开。

1. ハリーはまた望遠鏡に目を当て、焦点を合わせ直して、今度は金星を観測した。星座図を見下ろし、金星をそこに書き入れようとしたが、どうも何かが気になる。羊皮紙の上に羽根ペンをかざしたまま、ハリーは目を凝らして暗い校庭を見た。五つの人影が芝生を歩いているのが見えた。影が動いていなければ、そして月明かりがその頭を照らしていなければ、その姿は足元の芝生にのまれて見分けがつかなかっただろう。こんな距離からでも、ハリーはなぜか、集団を率いているらしい一番ずんぐりした姿の歩き方に見覚えがあった。

2. そこは長い豪華なホールが一番端で、黒っぽい木の床はピカピカに磨き上げられていた。ブルーの天井には金色に輝く記号がぞうがんされ、その記号が絶え間なく動き変化して、まるで空にかかった巨大な掲示板のようだった。両側の壁はピカピカの黒い木の腰板で覆われ、そこに金張りの暖炉がいくつも設置されていた。左側の暖炉からは、数秒ごとに魔法使いや魔女が柔らかいヒューツという音とともに現れ、右側には、暖炉ごとに出発を待つ短い列ができていた。

## 附录二：基频变化和音强变化M语言计算程序

```

clear;
clc;
filename='复件 sub 2 YYC f CN session 1 a L 001.csv';
data=CSVREAD(filename);           % 所要复制的原始数据 original data
fo=data(:,4);
fominibound=7;                   % 筛选范围 minimum
filter=fo>fominibound;
fo=fo(filter);
numoffo=numel(fo);
fodelta=zeros(numoffo-1,1);
for n=1:(numoffo-1)
    fodelta(n)=abs(fo(n+1)-fo(n));
end
fodeltamaxbound=20;             % ddt分析设定值
filter=fodelta<fodeltamaxbound;
fodelta=fodelta(filter);
fomean=mean(fo);                % 平均值结果 mean
fostd=std(fo);                  % 标准偏差 std
fodeltamean=mean(fodelta);      % ddt 分析结果 mean
R4=[fomean,fostd,fodeltamean];
power=data(:,3);
powerminibound=0;
filter=power>powerminibound;
power=power(filter);
numofpower=numel(power);
powerdelta=zeros(numofpower-1,1);
for n=1:(numofpower-1)
    powerdelta(n)=abs(power(n+1)-power(n));
end
powerdeltamaxbound=15;
filter=powerdelta<powerdeltamaxbound;
powerdelta=powerdelta(filter);
powermean=mean(power);          % 平均值结果 mean
powerstd=std(power);            % 标准偏差 std
powerdeltamean=mean(powerdelta); % ddt 分析结果 mean
R3=[powermean,powerstd,powerdeltamean];
conn = database('jpresult', ', ', ');
cols={'filename','powermean','powerstd','powerdeltamean','fomean',
'fostd','fodeltamean'};

```

```
resultforwrite={filename,powermean,powerstd,powerdeltamean,fomean,fostd,
fodeltamean};
insert(conn, 'jpresult', cols, resultforwrite);
close(conn);
```

本研究得到国家哲学社会科学基金重大项目(10&ZD126)、国家自然科学基金(30740040)、国家哲学社会科学基金(09CYY038)、江苏省高校自然科学研究重大基础研究项目(10KJA180051)、江苏省教育厅哲学社会科学项目(2012SJD740041)和南京航空航天大学青年科技创新基金(NR2011014)的资助,谨致谢忱。本研究为第一作者博士论文一部分,实验部分完成于日本北海道大学电子科学研究所。第一作者同时得到国家留学基金(201203070154)的资助。作者感谢邵纪东、Akira Toyomura在统计和技术上的协助。

## 参考文献

- 程工, 1999, 《语言共性论》。上海: 上海外语教育出版社。
- 高慧、周笃强、黄端生, 1999, 噪声对说话人语音的影响。《航天医学与医学工程》第1期, 72-75页。
- 韩纪庆、张磊、郑铁然, 2004, 《语音信号处理》。北京: 清华大学出版社。
- 河原英紀, 2003, 聴覚フィードバックの発声への影響——ヒトは自分の話声を聞いているのか。《日本音響学会誌》59,11: 670-675。
- 靳洪刚, 1997, 《语言获得理论研究》。北京: 中国社会科学出版社。
- 克里斯特尔, 沈家煊译, 2000, 《现代语言学词典》。北京: 商务印书馆。
- 李爱军, 2005, 语调研究中心理和声学等价单位。《声学技术》(西部声学会议)第3期, 13-17页。
- 卢华平、杨强等, 1996, 正常人汉语噪声竞争下言语测听的临床研究。《听力学及言语疾病杂志》第3期, 140-145页。
- 彭聃龄, 2004, 汉语信息加工及其认知神经机制的研究。《当代语言学》第4期, 302-320页。
- 索绪尔, 高名凯译, 2003, 《普通语言学教程》。北京: 商务印书馆。
- 田斌、易克初, 2003, 一种用于强噪声环境下语音识别的含噪Lombard及Loud语音补偿方法。《声学学报》第1期, 28-32页。
- 汪玲、郭德俊, 2000, 元认知的本质与要素。《心理学报》第32期, 458-463页。
- 王士元、彭刚, 2006, 《语言、语音与技术》。上海: 上海教育出版社。
- 杨亦鸣, 2007, 语言的理论假设与神经基础——以当前汉语的若干神经语言学研究为例。《语言科学》第2期, 60-83页。
- 赵元任著, 2006, 《赵元任语言学论文集》。北京: 商务印书馆, 382-399, 951-970页。
- 郑秋豫, 2008, 语篇韵律与上层讯息——兼论语音学研究方法与发现。《Language and Linguistics》, 第93期, 659-719页。
- 朱晓农, 2005, 《上海声调实验录》。上海: 上海教育出版社。
- 朱晓农, 2006, 《音韵研究》。北京: 商务印书馆。
- Balconi, M. 2008. Intentions and Communication: Cognitive Strategies, Metacognition and Social Cognition. Balconi M. (ed.) *Neuropsychology of Communication*, Springer-Verlag Italia: 159-172.

- Bond, Z. S. and Moore, T. J. et al. 1989. Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask. *The Journal of the Acoustical Society of America* 85(2): 907–912.
- Burnett, T. A. and Senner, J. E. et al. 1997. Voice F0 responses to pitch-shift auditory feedback: A preliminary study. *Journal of Voice* 11: 202–211.
- Burnett, T. A. and Freedl, M. B. et al. 1998. Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am*, 103(6): 3153–3161.
- Burnett, T. A. and Larson, C. R. 2002. Early pitch response is active in both steady and dynamic voice pitch control. *J. Acoustic Soc. Am* 112(3): 1058–1063.
- Burke, B. D. 1969. Reduced auditory feedback and stuttering, *Behav. Res. & Therapy* 7: 303–308.
- Carol, A. Fowler and Bruno, Galantucci. 2005. The relation of Speech perception and Production. In: David B. Pisoni and Robert E. Remez, eds., *The Handbook of Speech Perception*. Blackwell Publishing Ltd: 633–652.
- Carroll, D. W. 1994. *Psychology of Language* (2nd ed.). California: Cole Publishing Company, a division of Wadsworth, Inc., Belmont, 207–211.
- Castellanos, A. and Benedí, J. M. et al. 1996. An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard effect. *Speech Communication* 20: 23–35.
- Cerisara, C. and Rigazio, L. et al. 2004.  $\alpha$ -Jacobian environmental adaptation. *Speech Communication* 42: 25–41.
- Chao Y. 1964/2006. *Linguistic Essays*. Beijing: The Commercial Press, 784–791.
- Connell, B. 2002. Tone languages and the universality of intrinsic F0: evidence from Africa. *Journal of Phonetics* 30: 101–129.
- Dominey, P. F. and Dodane, C. 2004. Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics* 17: 121–145.
- Eden, G. F. and Moats, L. 2002. The role of neuroscience in the remediation of students with dyslexia. *Nature Neuroscience* (supplement) 5: 1080–1084.
- Elman, J. L. 1981. Effects of frequency-shifted feedback on the pitch of vocal productions. *J. Acoustic Soc. Am* 70(1): 45–50.
- Fernald, A. and Mazzei, C. 1991. Prosody and focus in speech to infants mothers' speech to new-borns. *Developmental Psychology* 27: 209–221.
- Flavell, J. H. 1979. Metacognition and cognitive monitoring: A new area of cognitive – developmental inquiry. *American Psychologist* 34: 906–911.
- Flavell, J. H. 1985. *Cognitive Development*. New Jersey: Englewood Cliffs, Prentice2 Hall Inc.
- Cruttenden, A. 2002. *Intonation*. Peking and Cambridge: Peking University Press and Cambridge University Press.
- Guenther, F. H. 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review* 102: 594–621.
- Guenther, F. H. 2006. Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders* 39, Issue 5: 350–365.
- Gunji, A. and Hoshiyama, M. et al. 2000. Identification of auditory evoked potentials of one's own voice. *Clinical Neurophysiology* 111: 214–219.
- Gunji, A. and Hoshiyama, M. et al. 2001. Auditory response following vocalization: a magnetoencephalographic Study. *Clinical Neurophysiology* 112: 514–520.
- Hansen, J. H. L. 1996. Analysis and compensation of speech under stress and noise for environment robustness in speech recognition. *Speech Communication* 20: 151–173.

- Hauser, M. D. and Chomsky, N. et al. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298: 1559–1579.
- Hockett, C. 1966. *A Course in Modern Linguistics*. New York: Macmillan.
- Hirano, S. and Kojima, H. et al. 1997. Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport* 8: 2379–2382.
- Jones, J. A. and Munhall, K. G. 2000. Perceptual calibration of F0 Production: Evidence from feedback perturbation, *J. Acoustic Soc. Am* 108: 1246–1251.
- Jones, J. A. and Munhall, K. G. 2002. The role of auditory feedback during phonation: studies of Mandarin tone Production. *Journal of Phonetics* 30: 303–320.
- Jones, J. A. and Munhall, K.G. 2003. Learning to produce speech with an altered vocal tract: The role of auditory feedback. *J. Acoust. Soc. Am* 113: 532–542.
- Jones, J. A. and Munhall, K.G. 2005. Remapping auditory-motor representations in Voice Production. *Current Biology* 15: 1768–1772.
- Junqua, J. C. 1993. The Lombard reflex and its role on human listeners and automatic speech recognizers. *The Journal of Acoustical Society of America* 93(1): 510–524.
- Junqua, J. C. 1996. The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. *Speech Communication* 20: 13–22.
- Kalinowski, J. and Armson, J. 1993. The effects of alterations in auditory feedback on stuttering frequency. *Language and Speech* 36: 1–16.
- Kawahara, H. and Kato, H. 1996. Effects of Auditory Feedback on F0 Trajectory Generation. The 4th International Conference on Spoken Language Processing. Philadelphia: PA, USA, October 3–6, ISCA Archive. [http://www.isca-speech.org/archive/icslp\\_1996](http://www.isca-speech.org/archive/icslp_1996).
- Keith, J. 2005. Speaker Normalization in Speech Perception. In: David B. Pisoni and Robert E. Remez eds., *The Handbook of Speech Perception*. Blackwell Publishing Ltd: 363–389.
- Kuhl, P. K. and Andruski, J. E. et al. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277: 684–686.
- Lee, B. S. 1950. Effects of delayed speech feedback. *J. Acoust. Soc. Am* 22: 824–826.
- Lincoln, M. and Packman, A. et al. 2006. Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders* 31: 71–89.
- Max, L. and Gauenthaler, F. H. et al. 2004. Unstable or insufficiently activated internal models and feed-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders* 31: 105–122.
- McGuire, P. K. and Silbersweig, D. A. et al. 1996. Functional neuroanatomy of verbal self-monitoring. *Brain* 119: 907–917.
- Moore, B. C. J. 2004. Unstable or insufficiently activated internal models and feed-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders* 31: 105–122.
- Muellerleile, S. 1981. Portable Delayed Auditory Feedback Device: A Preliminary Report. *Journal of Fluency Disorders* 6: 361–363.
- NATKE, U. and GrosseKarl, J. et al. 2001. Fluency, fundamental frequency, and speech rate under frequency-shifted auditory feedback in stuttering and nonstuttering persons. *Journal of Fluency Disorders* 26: 227–241.
- Packman, Ann. 2012. Theory and therapy in stuttering: A complex relationship. *Journal of Fluency Disorders* 37(4): 225–233.
- Paulesu, E. and Démonet, J. F. 2001. Dyslexia: Cultural Diversity and Biological Unity. *Science* 291: 2165–2167.

- Price, C. J. and Wise, R. J. 1996. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain* 119: 919–931.
- Saberi, K. and Perrott, D. R. 1999. Cognitive restoration of reversed speech. *Nature* 398: 760.
- Siok, W. T. and Perfetti, C. A. et al. 2004. Biological abnormality of impaired reading is constrained by culture. *Nature* 431: 71–76.
- Skowronski, M. D. and Harris, J. G. 2006. Applied principles of clear and Lombard speech for automated intelligibility enhancement in noisy environments. *Speech Communication* 48: 549–558.
- Snow, C. E. and Ferguson, C. A. 1997. *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press.
- Sternberg, R. J. 1994. *Encyclopedia of Human Intelligence*. London: Macmillan Publishing House.
- Steinberg, D. D. and Nagata, H. et al. 2002. Psycholinguistics: Language, mind and world. London: Longman. Reviewed by Natalie Sciarini-Gourianova, Guilford, CT, *Language* 78: 773–775.
- Studebaker, G. A. and Taylor, R. et al. 1994. The effect of noise spectrum on speech recognition performance-intensity functions. *Journal of Speech and Hearing Research* 37(2): 439–448.
- Toyomura, A. and Koyama, S. et al. 2007. Neural correlates of auditory feedback control in human. *Neuroscience*, 146: 499–503.
- Tremblay, S., Shiller, D. M. and Ostry, D. J. 2003. Somatosensory basis of speech Production, *Nature* 423, 866–869.
- Whalen, D. H. and Levitt, A. G. 1995. The universality of intrinsic F0 of vowels. *Journal of Phonetics* 23: 349–366.
- Wassink, A. B. and Wright, R. A. et al. 2007. Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics* 35, Issue 3: 363–379.
- Wildgruber, D. and Ackermann, H. et al. 1996. Functional lateralization of speech production at primary motor cortex: a fMRI study. *Neuroreport* 7 (15–17): 2791–2795.
- Wolpert, D. M. and Ghahramani, Z. et al. 1995. An internal model for sensorimotor integration. *Science* 269: 1880–1882.
- Wolpert, D. M. and Miall, R. C. 1996. Forward Models for Physiological Motor Control. *Neural Networks* 9: 1265–1279.
- Xu, Yisheng and Gandour, J. T. et al. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *J. Acoust. Soc. Am* 120(2): 1063–1073.
- Yates, A. J. 1963. Delayed Auditory Feedback. *Psychological Bulletin* 60: 213–232.
- Zhao, Y. and Jurafsky, D. 2009. The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics* 37: 231–247.

## Bionotes

*Xu Yizhong* is an associated professor in the College of Foreign Languages, Nanjing University of Aeronautics and Astronautics. He has a PhD in Linguistics and Applied Linguistics from Zhejiang University. His research area is cognitive linguistics, experimental phonetics and comparative linguistics. E-mail: Calebxuxu@163.com

*Yang Yiming* is a professor at the Jiangsu Key Laboratory of Language and Cognitive Neuroscience, Jiangsu Normal University. His research area is neurolinguistics and theoretical linguistics. He is the director of the Jiangsu Collaborative Innovation Center for Language Ability, Jiangsu Normal University, China. E-mail: yangym@jsnu.edu.cn