

《心理学报》审稿意见与作者回应

题目：老龄化对 F0 轮廓线索在汉语言语识别去掩蔽作用的影响

作者：吴梅红

第一轮

审稿人 1 意见：

本文考察了老龄化对 F0 轮廓线索在汉语言语识别中去掩蔽作用的影响，研究发现老年人在言语掩蔽下不能从动态 F0 轮廓线索中受益。文章在综述、方法、结果分析和讨论中均存在较多问题，有以下建议供作者参考。

回应：

非常感谢审稿专家的意见，已经根据审稿专家的建议对每个部分的问题进行了修改。在以下的意见回应中，将根据审稿专家对每个部分的意见进行一一答复，具体修改及回复内容如下所示。

意见 1：文献综述

文献综述中有关为什么要研究老年人在汉语言语感知中对 F0 轮廓信息的加工方面的阐述比较泛化。建议作者进一步梳理相关影响因素，结合具体研究结果进行细致地阐述。

第一，根据文献综述中的阐述，老年人在背景掩蔽的情况下难以识别言语的原因包括了：1) 老年人利用声音线索感知分离语音流的能力较差，2) 外周听觉功能衰退、中枢听觉处理下降以及认知能力的变化。作者需要引用文献，深入阐述这些能力的变化，对于 F0 轮廓线索的知觉可能产生哪些具体的影响。

回应：

感谢审稿专家的意见。已经根据专家的建议重写了文献综述部分。针对审稿专家的意见，对前言部分进行了相应的修改。具体修改如下所示：

文献综述中从以下几点思路展开阐述研究老年人在汉语言语感知中对 F0 轮廓信息加工的原因：

1)基频(F0)轮廓对日常交流中的言语识别非常重要,可以作为从背景噪声中分离语音的感知线索。相关论述补充如下(请详见引言 1 部分第 2 段, P1):

“嘈杂环境下听者会利用某些与目标言语特征有关联的知觉线索来促进对被掩蔽的目标言语的识别(李量等, 2017)。基频(F0)轮廓对嘈杂环境下的言语感知非常重要,它的功能不仅是强调特定的单词和标记短语边界(Cutler, Dahan, & vanDonselaar, 1997),还可以作为从掩蔽声音中分离语音的感知线索(Binns & Culling, 2007)。近年来许多研究者研究了 F0 轮廓线索对嘈杂环境下言语感知的影响(Binns & Culling, 2007; Chen, Yang, Wu, & Moore, 2018; Miller, Schlauch, & Watson, 2010; Patel, Xu, & Wang, 2010; Shen & Souza, 2019; Wu, 2019),一致认为 F0 轮廓动态变化对噪声下的言语可懂度具有重要作用。”

2)动态基频(F0)轮廓线索对年轻人在噪声下言语识别的作用已经在多个数据集上得到一致验证,但对老年人的影响还知之甚少。

相关论述补充如下(请详见引言部分 P2 第 4 段):

“动态 F0 轮廓线索有助于年轻听者在言语掩蔽下有效的语音识别已经在多个数据集上得到了一致验证,然而关于老年听者的研究还很欠缺。探讨老年人能够利

用哪些知觉线索以减少他们在嘈杂环境中的言语识别困难具有重要意义。在多人同时说话的嘈杂环境中老年听者是否能从动态 F0 轮廓线索中获益,目前尚不清楚。”

3)对于听力正常的年轻听者来说,基频(F0)轮廓的动态变化有助于言语掩蔽下有效的语音识别。然而 F0 动态变化是否可以帮助听力正常的老年听者从掩蔽言语流中分离出目标,还有待阐明。

其中,对老年人在背景掩蔽情况下面临更大言语识别困难原因,以及关于年龄相关性听力变化对于老年人感知和利用 F0 轮廓线索可能产生哪些影响的阐述,我们在论文引言部分的第 2 页第 5 段进行了补充(请详见论文引言部分,P2-P3),具体如下:

“相对于听力正常的年轻人而言,动态 F0 线索对老年听者在噪声下言语感知的影响更为复杂和多变。老年人对 F0 轮廓线索的感知和利用可能会比年轻人更加困难。由于年龄增加所带来的外周听觉系统功能的下降,比如听觉阈限的提高、听觉滤波器带宽的增加、时间分辨率的下降等均可导致老年人的言语加工能力的降低。因此,老年人感知动态 F0 线索的能力可能会因各种阈上听力缺陷而下降,这其中包括时间编码不良(Grose & Mamo, 2010; Hopkins & Moore, 2011)、频率选择性降低(Hopkins & Moore, 2011),以及频率调制的退化神经表征(Clinard & Cotter, 2015)。这种与年龄相关的听觉时间加工缺陷会影响老年人对语音中时变的基频准确地进行编码,从而影响老年人感知言语中 F0 轮廓线索的能力。

除此之外,老年人高级认知加工功能的下降也会引起他们在嘈杂环境下的言语识别困难 (e.g., Tun, McCoy, & Wingfield, 2009; Wingfield & Tun, 2001; Humes, Lee, & Coughlin, 2006)。值得注意的是,纯音听觉阈值并不能完全直接反映老年人的听觉能力 (Ruggles, Bharadwaj, & Shinn-Cunningham, 2011)。例如,从感音神经性听力损失和听力正常的老年人中收集的频率辨别数据表明,在听力图相似的老年人中感知音高线索的能力差异很大(Moore & Peters, 1992)。此外,即使拥有临床正常的纯音听觉阈值,老年言语感知困难仍可能发生(Fullgrabe, Moore, & Stone, 2014)。与年龄相关的噪声中言语理解差异可能是由于与纯音听阈升高无关的其他听觉能力的下降,并涉及中枢听觉或认知处理过程(Fullgrabe, Moore, & Stone, 2014)。

认知水平上,老年人在嘈杂环境下处理和理解言语的能力会受到选择性注意、工作记忆、抑制性控制和认知加工速度普遍减慢的影响(e.g., van Knijff, Coene, & Govaerts, 2018; Schneider, Avivi-Reich, & Daneman, 2016)。相对于年轻人,老年人可能需要更多的认知资源,以便在背景噪声存在的情况下解析言语信号,尤其是当输入信号由于听觉时间处理缺陷而进一步消退时,这对自上而下的处理提出了更高的要求。因此,老年人在多人说话的嘈杂环境下处理 F0 轮廓信息会比年轻人更具有挑战性。”

4)虽然许多研究已经对非声调语言(e.g. 英语)中 F0 轮廓线索对噪声下言语感知的作用进行了探讨,但是对于像汉语这样的声调语言来说,噪声环境下言语感知中 F0 轮廓信息是如何处理的却鲜有涉及。关于老年听者的研究更为欠缺。同英语相比,汉语言不论是在其自身特性上还是在抗掩蔽特征上都有很大的差异(Kang, 1998)。因此,汉语语境下 F0 轮廓信息在言语感知中是如何处理的需要得到更多关注。我们在论文引言部分的第 3 页第 5 段进行了补充(请详见论文引言 1 部分,P3-P4),具体如下:

“噪声下言语感知能力下降会限制老年人人际交流质量,进而影响到老年人的身心健康。老年人在噪声下言语感知困难背后的机制仍不清楚,治疗方案也很少。动态 F0 轮廓线索对日常交流中的言语识别很重要,听觉加工老化是否会影响老年人对动态 F0 轮廓线索的感知和利用,现有的数据集还很有限。F0 轮廓动态变化是否能够帮助老年人从掩蔽中分离出目标言语还有待阐明。嘈杂环境下老年人在汉语言语感知中对 F0 轮廓信息是如何处理的更是知之甚少。值得注意的是,汉语单字的声调具有语义意义,

其基频的波动较英语的活跃，这有助于抵抗信息掩蔽。另外，汉语的清辅音较英语的多，会更加受噪音的掩蔽(Kang, 1998)。因此，汉语语境下 F0 轮廓信息在言语感知中是如何处理的需要得到更多关注。本文将基于汉语成句语音，尝试阐明 F0 轮廓线索的操作对言语掩蔽下的汉语言语识别去掩蔽作用受老龄化的影响。”

从以上几点论述出发，本文评估了老年人与年轻人在言语噪声环境下聆听具有自然动态 F0 轮廓与对 F0 轮廓操作调节后的汉语语句的言语识别能力，探讨老龄化对 F0 轮廓线索在言语掩蔽下汉语言语识别中作用的影响。

引言部分新增加的参考文献如下所示：

- Cutler, A., Dahan, D., & vanDonselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141-201.
- Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *Journal of the Acoustical Society of America*, 120(5), 2926-2937.
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G.(2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37), 15516–15521.
- Moore, B. , & Peters, R. W. . (1992). Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. *Journal of the Acoustical Society of America*, 91(5), 2881.
- Fullgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci*, 6, 347.
- van Knijff, E. C., Coene, M., & Govaerts, P. J. (2018). Speech understanding in noise in elderly adults: the effect of inhibitory control and syntactic complexity. *Int J Lang Commun Disord*, 53(3), 628-642.
- Wright, H. H. (2016). *Cognition, language and aging*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research*, 106(1-2), 95-104.
- Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *J. Acoust. Soc. Amer.*, 123, 489–499.
- Singh, G., Pichora-Fuller, M. K., & Schneider, B. A. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *Journal of the Acoustical Society of America*, 124(2), 1294-1305.
- Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *J Am Acad Audiol*, 18, 548–558.
- Bernstein JG, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2009 May;125(5):3358-72.
- Wang, J., Hua, S., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in mandarin chinese speech intelligibility. *The Journal of the Acoustical Society of America*, 134(1), EL91-7.
- Xu L., Zhou N. (2011). Tonal languages and cochlear implants. In: Zeng F.G., Popper A.N., Fay R.R., editors. *Auditory Prostheses: New Horizons*. Springer; New York. pp. 341–364.
- Darwin, C. J., Listening to speech in the presence of other sounds. *Phil. Trans. R. Soc. B*, 2008, 363: 1011–1021.
- Kang, J. (1998). Comparison of speech intelligibility between english and chinese. *The Journal of the Acoustical Society of America*, 103(2), 1213-1216.

意见 2:

第二，研究中有关 F0 轮廓的变化包括了自然的、扁平的和拉伸的 F0 轮廓条件。已有研究对于这三种条件的对比所发现的具体结果是什么？在此基础上提出本研究所关注的问题及其贡献。

回应:

感谢审稿专家的意见。根据专家的意见已经在引言部分增加了对 F0 轮廓线索的操作对噪声下言语识别影响的更为详细具体的描述。

我们在论文引言部分第 3 段开始进行了补充(详见论文引言 1 部分, P1-P2),具体修改内容如下:

“基频(F0)轮廓变化有助于多语流中的说话人分离，对日常交流中的言语识别非常重要(Greenberg, Ainsworth, Popper, & Fay, 2004)。作为不利声学环境下促进言语识别的强有力线索之一，F0 轮廓动态变化能够增强语音流的感知连续性(Dannenbring, 1976)，并提高对背景掩蔽下的目标言语识别(Shen & Souza, 2019)。因此，许多研究都将兴趣集中在 F0 轮廓线索对言语感知的贡献上(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。研究者通过对自然语句 F0 轮廓的调节操作，例如压扁(Binns & Culling, 2007)、拉伸(Miller et al., 2010)或反转(e.g., Hillenbrand, 2003)基频 F0 轮廓，来检验对不同 F0 轮廓调节后对言语可懂度的影响。

对于英语语音的感知，以往的研究结果显示，无论是在安静状态下还是在稳定的语谱噪声中，通过扁平化英语句子的 F0 轮廓，语音可懂度会降低(e.g., Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。在对这些发现提出的解释中认为 F0 轮廓的上升和下降会将听者的注意力引向言语的内容词；因此，如果没有这些线索，言语的可理解性就会下降(Binns & Culling, 2007)。Miller 等人(2010)也证明了被调节的 F0 轮廓对噪声下的语音识别有害，拉伸的 F0 轮廓与保留自然动态 F0 轮廓相比并没有发现更有利的效果。Hillenbrand(2003)也发现自然动态 F0 轮廓线索对于不利条件下的语音识别更有帮助，语音中的平坦或倒转 F0 轮廓会导致言语可懂度的下降。这些研究结果一致表明自然动态变化的 F0 轮廓有利于英语语音的识别。

基频 F0 轮廓线索对噪声环境下的英语语音感知作用已经得到了广泛的研究(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。对于汉语语音感知的研究，现有的数据集还很有限，但动态 F0 线索有助于听者在背景噪声下言语识别的这种效应在汉语语言中也得到了证实。Patel 等人(2010)使用有意义的汉语句子为目标刺激，证明了在安静的环境中，具有平坦 F0 轮廓的汉语句子与具有自然 F0 模式的句子一样容易理解(Patel et al., 2010)；然而，正如其他几项研究(e.g., Wang et al., 2013; Wu, 2019)报告的那样，在稳态噪声或嘈杂语掩蔽下，F0 轮廓线索被调节的句子的言语可懂度较低。

在汉语普通话中，F0 轮廓是单音节水平上声调感知的主要线索(Xu & Zhou, 2011)，另外它还与句子语调和音高的感知有关。Wu(2019) 使用中文无意义句子作为言语刺激，对句子的语调进行扁平化或拉伸的调节操作，同时保留每个单音节的声调信息。与英语语音感知的研究(Binns & Culling, 2007; Miller et al., 2010)结果一致，具有自然 F0 轮廓线索的语音比对 F0 轮廓线索调节操作后的语音更容易被感知。此外，Wu(2019)的研究还表明压扁或拉伸 F0 轮廓线索会对噪声环境下的言语可懂度产生不利影响，尤其是在多人说话的言语掩蔽下。这些结论说明了动态 F0 轮廓线索有助于言语掩蔽下的汉语言语识别。F0 轮廓线索在汉语言语识别去掩蔽中具有重要作用。”

本部分内容涉及的参考文献如下所示：

- Greenberg, S., Ainsworth, W., Popper, A., & Fay, R. (2004). *Speech processing in auditory system*: Springer Press.
- Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, 30(2), 99-114.
- Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *Journal of the Acoustical Society of America*, 122(3), 1765-1776.
- Miller, S. E., Schlauch, R. S., & Watson, P. J. (2010). The effects of fundamental frequency contour manipulations on speech intelligibility in background noise. *Journal of the Acoustical Society of America*, 128(1), 435-443.
- Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility. *Journal of the Acoustical Society of America*, 114(4), 2338.
- Shen, J., & Souza, P. E. (2019). The ability to glimpse dynamic pitch in noise by younger and older listeners. *Journal of the Acoustical Society of America*, 146(3), E1232-E1237.
- Wang, J., Hua, S., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in mandarin chinese speech intelligibility. *The Journal of the Acoustical Society of America*, 134(1), E191-7.
- Wu, M. H. (2019). Effect of F0 contour on perception of Mandarin Chinese speech against masking. *Plos One*, 14(1).
- Xu L., Zhou N. (2011). Tonal languages and cochlear implants. In: Zeng F.G., Popper A.N., Fay R.R., editors. *Auditory Prostheses: New Horizons*. Springer; New York. pp. 341-364.

意见 3:

第三，文献引用错误。吴翰林等（2020）的研究中并未涉及到“自上而下的加工补偿以及听觉信息缺失的问题”。

回应:

谢谢专家的指出，该文献在这边进行引用的确会引起读者误解。由于重新写的引言中不再涉及到“老年人对晶体智力的保持能力”的相关阐述，已将原文中的这整句话“另外，老年被试还可以利用其保留完好的晶体智力(吴翰林等人,2020)通过自上而下的加工补偿由于听觉衰退引起程中信息缺失(Kricos, 2006)。”及其中的文献引用都进行了删除。

意见 4: 方法部分

第一, 青年组和老年组被试各自仅有 12 人, 建议增加被试。或者要用恰当的统计指标表明当前的被试量是足够的。

回应:

感谢专家的建议。我们使用 G*Power3.19 软件进行先验分析, 估计出所需要的样本量, 应用统计指标验算本文中的数据量能否大于估计的样本量, 以表明当前的被试量是足够的。

具体方法: 选择 G*power 中最常见的统计力分析(power analysis)类型中的先验(priori)分析, 即根据所给的 α 水平, 统计力水平(1- β)和效应大小(H0 和 H1 的差异), 计算样本容量。

分析结果显示需要每组至少有 10 个样本(Total sample size:20), 表明本文中当前的被试量(青年组和老年组, 每组 12 个)是足够的。

意见 5:

第二, 无意义的语句材料是否经过了意义的评定, 相关语义等级的结果需要报告。

回应:

感谢专家的建议。本文实验中用到的目标言语刺激为中文无意义句子。所有的句子均来自于北京大学言语听觉研究中心自建的语料库, 并已被许多研究者 (Chen et al., 2018; Wu, 2019; Wu et al., 2012; Yang et al., 2007)使用。为了确保这些句子都是无意义的, 每个句子中所用的两个名词和动词在所选的 9 年的《人民日报》中共同出现的概率都做了严格限制。

意见 6:

第三, 目标语句的为中文无意义语句, 被试的任务是复述听到的句子, 作者设置关键词的目的是什么? 文中并未做出清楚的交待。

回应:

感谢专家的意见。根据标准的言语测听实验中检查被试的言语识别能力所用的言语识别率测试方法, 设置关键词的目的是为了测量被试的言语识别率。

对实验中具体测试方法描述的解释如下:

实验测试中用到的每个目标句子均包含三个关键成分, 主语、谓语和宾语。这三种成分也是三个关键的双字词。实验中被试的任务是在一次测试(trial)的所有声音结束之后尽可能多地重复所听到的目标句子。实验人员记录被试对每个句子中三个关键词的复述, 每个关键词的两个音节均复述正确则计 2 分。

根据专家的建议, 已经在文中增补了本项实验在言语测听测试中是对中文无意义语句中关键词识别率测量的描述, 具体修改如下所示:

1)在第 5 页第 3 段的方法部分增补了对语句中设置关键词的强调表述, 具体如下(详见论文实验方法部分 2.3 刺激材料章节, P5):

“每句中文无意义语句都包含主语(名词)-谓语(动词)-宾语(名词)结构。例如, 在一个样例句子“他的磐石可能比赛这个员工”中, 下划线的主、谓、宾语为句子的三个关键词, 每个关键词都是一个双音节词。该无意义语句本身的句子结构对关键词的识别不提供任何上下文信息。”

2)在论文第 7 页实验结果部分第 3 段增加了以下内容(详见论文实验结果部分, P7):

“实验过程中对三个关键词进行计分, 每正确复述一个音节计一分, 一个句子满分为六分。之后计算每名被试在每种实验条件下(F0 轮廓类型×信噪比)的识别率以用于统计分析。”

意见 7: 实验结果

第一, 对图 3 的描述建议删除, 建议直接呈现数据统计的结果。

回应:

感谢专家的意见。

已经根据专家的建议将图 3 的描述删除, 直接呈现数据统计的结果。

意见 8:

第二, 对于目标句的识别率是如何计算的? 是根据关键词还是整个句子? 如何测量整个句子的识别率?

回应:

本文中的言语识别率计算方法, 与许多研究者(Huang et al., 2010; Wu et al., 2012; Lu et al., 2018; Cao et al., 2011; Wu, 2019)论文中的言语识别率统计方法相同, 实验测试的是整个句子中三个关键词的复述正确率。

根据专家的意见, 本文在方法部分对言语识别率的测试方法做了描述补充, 在论文第 7 页实验结果部分增加了以下内容:

“在实验过程中我们对三个关键词进行计分, 每正确复述一个音节计一分, 一个句子满分为六分。之后我们计算每名被试在每种实验条件下(F0 轮廓类型×信噪比)的识别率以用于随后的统计分析。”

本部分内容涉及的参考文献如下所示:

Huang, Y., Xu, L., Wu, X., & Li, L. (2010). The effect of voice cuing on releasing speech from informational masking disappears in older adults. *Ear and Hearing*, 31(4), 579-583.

Wu, M., Li, H., Hong, Z., Xian, X., Li, J., & Wu, X., et al. (2012). Effects of aging on the ability to benefit from prior knowledge of message content in masked speech recognition. *Speech Communication*, 54(4), 529-542.

Chao, W., Cao, S., Zhou, F., Wang, C., Wu, X., & Li, L. (2012). Masking of speech in people with first-episode schizophrenia and people with chronic schizophrenia. *Schizophrenia Research*, 134(1), 33-41.

Lu, L., Bao, X., Chen, J., Qu, T., Wu, X., & Li, L. (2018). Emotionally conditioning the target-speech voice enhances recognition of the target speech under "cocktail-party" listening conditions. *Attention Perception & Psychophysics*.

Cao, S., Li, L., & Wu, X. (2011). Improvement of intelligibility of ideal binary-masked noisy speech by adding background noise. *The Journal of the Acoustical Society of America*, 129(4), 2227.

Wu (2019). Effect of F0 contour on perception of Mandarin Chinese speech against masking. *Plos One*, 14(1).

意见 9:

第三, 建议用混合线性模型对被试的识别做分析, 同时将被试和刺激作为随机变量。

回应:

感谢专家的建议。使用混合线性模型分析数据具有许多优点: 线性混合模型在计算过程中采用最大随机效应结构, 将被试和刺激定义为交叉随机效应同时纳入模型进行计算, 使得结果更为稳定和统一。而且, 线性混合效应模型能将所有的原始数据纳入到模型中进行统计, 无需对数据进行平均后再作比较, 其数据利用率更高。

根据专家的建议, 我们采用线性混合效应模型对被试的识别数据进行了重新统计分析, 具体分析结果如下所示:

“实验使用 SPSS 统计软件, 采用线性混合效应模型(linear mixed-effects models, LMMs; Baayen et al., 2008)对对被试的识别阈值进行统计分析。在本研究的模型中, 包括 F0 轮廓类型条件(压扁、自然、拉伸)、年龄组(年轻、老年)的固定效应, 同时考虑被试随机和项目随机效应。

实验(固定效应检验)结果显示, F0 轮廓类型($F(2,44)=7.14, p=0.002$)和年龄组组间差异($F(1,22)=73.21, p < 0.001$)对被试的识别成绩均有显著影响, F0 线索类型和年龄组的交互作用显著($F(2,44)=4.93, p=0.012$)。固定效应参数估计进一步显示, 老年被试组的言语识别表现明显差于年轻被试组($\beta=-7.73E0, SE=0.903553, t=-8.557, p < 0.001$)。

来自年轻被试组的数据分析显示, 在对 F0 线索操作的影响方面, 不同 F0 轮廓类型对被试识别成绩的固定效应具有统计学意义, $F(2,22)=6.874, p=0.005$ 。在模型中将自然 F0 轮廓条件作为基线条件, 对 F0 轮廓被调节的条件和自然 F0 轮廓条件之间的识别成绩进行比较。结果显示, 与自然动态 F0 轮廓线索条件相比, 压扁 F0 轮廓条件下的言语识别表现明显较差($\beta=2.9158E0, SE=0.876849, t=3.325, p=0.003$); 拉伸 F0 轮廓条件

下的言语识别成绩也比自然动态 F0 轮廓条件下的显著降低($\beta=2.7033E0$, $SE=0.876849$, $t=3.083$, $p=0.005$)。

对老年被试组的识别成绩分析表明,不同 F0 轮廓线索类型对识别成绩的固定效应无统计学意义, $F(2,22)=0.498$, $p=0.614$ 。压扁 F0 轮廓条件($\beta=0.339760$, $SE=0.340560$, $t=0.998$, $p=0.329$)或拉伸 F0 轮廓条件($\beta=0.163173$, $SE=0.340560$, $t=0.479$, $p=0.637$)下的言语识别成绩与自然动态 F0 轮廓线索条件下的差异均不显著。”

本部分内容涉及的参考文献如下所示:

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
张文彤, & 董伟. (2013). *SPSS 统计分析高级教程*. 高等教育出版社.

意见 10:

第四, p 值的呈现不符合《心理学报》的规定; 统计量是否斜体需要仔细对照学报要求进行修改。

回应:

感谢专家的意见, 已经严格按照学报的要求进行了相应的修改。

意见 11:

第五, 青年组和老年组的纯听阈限不同, 实验中需要考虑这一因素对结果的影响。

回应:

感谢专家的意见。

老年人在噪声中理解言语的困难程度差异很大, 并且不能很好地从纯音听觉阈限图中预测出来 (Fullgrabe, Moore, & Stone, 2014)。对于纯音听觉阈值并不能完全反映老年人的听觉能力以及他们噪声下的言语识别困难。我们在第 3 页第 2 段的引言部分增加了相关内容做了补充解释, 具体如下:

“值得注意的是, 纯音听觉阈值并不能完全直接反映老年人的听觉能力(Ruggles, Bharadwaj, & Shinn-Cunningham, 2011)。例如, 从感音神经性听力损失和听力正常的老年人中收集的频率辨别数据表明, 在听力图相似的老年人中感知音高线索的能力差异很大(Moore&Peters,1992)。此外, 即使拥有临床正常的纯音听觉阈限值, 老龄言语感知困难仍可能发生(Fullgrabe, Moore, & Stone, 2014)。与年龄相关的噪声中言语理解差异可能是由于与纯音听阈升高无关的其他听觉能力的下降, 并涉及中枢听觉或认知处理过程(Fullgrabe et al., 2014)。”

本部分内容涉及的参考文献如下所示:

Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G.(2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37), 15516–15521.
Fullgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci*, 6, 347.
Moore, B., & Peters, R. W. (1992). Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. *Journal of the Acoustical Society of America*, 91(5), 2881.

意见 12: 讨论

第一, 在讨论部分, 作者指出: “老龄化会降低汉语普通话语句基频轮廓感知的范畴性, 并且降低了用于区分 F0 轮廓变化的心理声学能力。”对此, 我有两点疑问: 1) 基频轮廓的范畴性在本文的研究中是如何体现的? 如何比较青年组和老年组轮廓感知的范畴性? 2) 什么是心理声学能力? 用于区分 F0 轮廓变化的心理声学能力指的是什么?

回应:

感谢专家的意见, 我们重新思考了这句总结话语的表述, 认为在这里堆砌专业术语并不合适。为了避免表述不严谨且晦涩难懂, 我们重写了整个讨论部分, 并已经删除了这句总结的语句。

意见 13:

第二, 本文主要研究目的是 F0 轮廓变化对语言感知中所出现的老龄化现象, 讨论要围绕“老龄化”主题, 不要分别讨论青年人和老年人的结果模式。

回应:

非常感谢专家的意见。根据专家的建议, 我们在讨论部分把子标题 4.1(动态 F0 轮廓线索有助于年轻人言语掩蔽下的汉语语句识别)和子标题 4.2(扁平化或拉伸语句 F0 轮廓对老年人言语掩蔽下目标汉语识别的影响)下分别讨论青年人和老年人结果模式的内容合并成一个段落并对所讨论的内容作了相应凝练。

我们将上一轮修改论文中子标题 4.1 和 4.2 下的全部内容进行了修改, 凝练为本论文中讨论章节的前三段讨论(请详见讨论部分 P8,第 4-6 段), 并修改了 F0 轮廓变化对言语感知中所出现的老龄化现象的总结表述, 具体如下(P8,第 6 段):

“在复杂听觉场景中, 听觉系统可以利用声音信号自身的线索—非高层的语言知识, 对混合的声音流进行分流和重组(Bregman, 1990)。作为不利听觉条件下辅助言语感知有益的声学线索, 基频轮廓动态变化能够促进对语音流感知连续性的增强(e.g., Dannenbring, 1976), 这有助于听者更好地将目标语音与背景噪声分离。本文中的数据结果也证明了动态 F0 轮廓线索在言语感知去掩蔽中的重要作用, 研究结果支持了许多研究者(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019; Wu, 2019)关于动态 F0 轮廓线索有助于年轻听者在言语掩蔽下目标言语识别的一致结论。本文的数据结果还表明, 老龄化会降低对动态 F0 轮廓线索的感知和利用, 使得老年听者难以利用 F0 轮廓线索来促进掩蔽下的目标言语识别。实验结果发现对于老年人来说, F0 轮廓线索对促进他们在目标语和掩蔽语之间知觉分离作用变得有限, 说明在多人说话的言语掩蔽下老年人很可能感知并追踪 F0 轮廓动态变化出现困难。”

意见 14:

第三, 领域一般性的认知能力和领域特异性的语音知觉加工能力都会影响老年人的表现, 研究者需要对此做出充分讨论。

回应:

感谢专家的意见。已经根据专家的意见重新写了讨论章节的内容, 具体修改如下:

1)在第 9 页讨论章节第二段对老龄化会降低老年人对动态 F0 轮廓线索的感知做了补充讨论, 具体的修改部分如下所示:

“在多人同时说话的嘈杂环境下, 听者会在知觉层次上利用一些线索来促进他们在目标语音和掩蔽语音之间的知觉分离, 提高对目标言语的有效识别(Schneider, Avivi-Reich, & Daneman, 2016)。嘈杂环境下年轻听者可以利用语音基频的差异来区分目标语音和掩蔽语音(Darwin, Brungart, & Simpson, 2003)。当目标语音和掩蔽语音同时存在时, 目标语与掩蔽语之间的 F0 差异能够加强听者对目标言语的跟踪, 进而提高言语识别率(Brungart, Simpson, Ericson, & Scott, 2001)。然而老龄化会对处理竞争声音之间的基频差异的能力造成有害影响。有研究证据显示老年人在处理安静环境中声音的基频信息时有更大的困难(Lee, 2013)。老年被试的 F0 辨别阈值明显大于年轻被试(Souza, Arehart, Miller, & Muralimanohar, 2011; Vongpaisal & Pichora-Fuller, 2007); 老年人也表现出使用 F0 差异来感知竞争声的难度增加(e.g., Rossi-Katz & Arehart, 2009; Vongpaisal & Pichora-Fuller, 2007)。由于老年人在处理相互竞争的说话者之间的 F0 差异方面增加了难度, 因此 F0 处理的缺陷可能会导致老年人在嘈杂环境下从 F0 线索中获益的能力存在年龄相关性下降。本文的研究结果也表明了多人说话的言语掩蔽下老年人遵循 F0 轮廓动态变化变得困难。”

2)在第 9 页第四段到第 10 页的第二段展开讨论了老龄化会对老年人受益于 F0 轮廓

线索的能力产生不利影响，修改后的内容具体如下所示：

“言语的时间结构分成为周期性线索、包络线索和精细结构信息三个不同的水平，这几种时域线索相互补充，传递了语音的完整信息(Rosen,1992)。给听力正常的听者播放言语语音时，F0感知的主导线索是分解的低频谐波 (Arehart, 1994; Oxenham, 2008)，即精细结构；此外 F0感知也可以从高频未分解谐波传递的高速包络波动中得出，即周期性(Oxenham, 2008)。许多证据表明老年人比年轻人在有言语掩蔽下的语音识别任务以及涉及周期性线索的心理物理任务方面表现更差(e.g.,Vongpaisal & Pichora-Fuller, 2007; Arehart et al., 2011)。这些研究中普遍存在的假设是，周期性线索有助于促进语音流分离，而老年人比年轻人更难意识到这一好处(Bologna et al.,2019)。已有的研究证明了在多人说话的言语识别任务中，为 F0 提供周期性线索可以提高句子识别能力(Stone et al., 2008)。然而老龄化会对周期性和精细结构知觉产生不利影响。例如，在脑干周期性线索的神经表征中观察到了年龄相关的下降(Clinard & Tremblay, 2013; Snyder & Alain, 2005)。老年人难以使用周期性线索来获取与说话者语音基本频率相关的信息(Souza et al., 2011)。因此，与年龄相关的周期性编码下降会对老年人受益于 F0 轮廓线索的能力产生不利影响。

此外，由于年龄相关的时域包络(e.g., Souza & Boike 2006; Souza et al. 2011)和精细结构(e.g. Vongpaisal & Pichora-Fuller 2007; Hopkins & Moore 2011; Smith, Delgutte, & Oxenham, 2002)线索处理方面的缺陷，老年人可能不能像年轻人一样使用和整合不同类型的线索进行言语识别。基频(F0)轮廓线索对老年人在目标言语识别中的贡献模式可能也不同于年轻人。

包络成分和精细结构成分信息对听觉感知具有不同的贡献。精细结构主要对语音的音调识别和声源定向起重要作用，而包络主要对言语的内容识别起重要作用(李量等 2017; Smith et al., 2012)。基频 F0 是一种精细结构的线索(Brown & Bacon, 2010)，有利于在噪声中的言语识别。然而老龄化会降低对精细结构的感知敏感度，从而影响对精细结构信息的编码能力(Pichora-Fuller, Schneider, Macdonald, Pass, & Brown, 2007)。这种对声音精细结构信息处理精确性的缺失会使得老年人在需要跟踪基频变化的情况下处于不利地位。因此，与年轻人相比，老年人在言语掩蔽下遵循 F0 轮廓动态变化更为困难。这种感知和跟踪 F0 轮廓动态变化的能力的年龄相关性下降，使得老年人利用 F0 轮廓线索促进他们将目标从背景掩蔽声中分离出来的作用变得有限。”

意见 15:

第四，作者根据对文献综述部分的建议修改后，对讨论部分也要重新做出全面修改。目前讨论部分的子标题过多，建议作者凝练主题，根据研究的主要目的进行讨论。

回应:

感谢专家的意见。已经根据专家的建议对讨论部分重新做出全面修改。

新修改后的讨论部分还删除了上一轮修改论文里子标题 4.3(老龄化降低了对动态基频线索感知和利用的心理声学能力)和子标题 4.4(在言语掩蔽下老龄化对从动态 F0 轮廓线索中获益的影响)下的所有的讨论内容。并根据研究的主要目的对“为什么动态 F0 轮廓线索对老年人言语掩蔽下的汉语语句识别没有促进作用”展开进一步的讨论。(详见修改后的论文中第四部分讨论章节,P8-P11)。

讨论部分增加的参考文献如下:

- Souza, P., Arehart, K., Miller, C.W., & Muralimanohar, R. K.. (2011). Effects of age on f0 discrimination and intonation perception in simulated electric and electroacoustic hearing. *Ear & Hearing*, 32.
- Vongpaisal, T., & Pichora-Fuller, M. K.. (2007). Effect of age on f0 difference limen and concurrent vowel identification. *Journal of Speech Language & Hearing Research*, 50(5), 1139.
- Rossi-Katz, J., & Arehart, K. H. (2009). Message and Talker Identification in Older Adults: Effects of Task, Distinctiveness of the Talkers' Voices, and Meaningfulness of the Competing Message. *Journal of Speech Language and Hearing Research*, 52(2), 435-453. doi:10.1044/1092-4388(2008/07-0243)
- Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition.

- Hearing Research, 235(1), 143–156.
- Brown, C. A., Tillery, K. H., Apoux, F., Doyle, N. M., & Bacon, S.P. (2015). Shifting fundamental frequency in simulated electric-acoustic listening: effects of f0 variation. *Ear & Hearing*, 37(1).
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci*, 336(1278), 367-373.
- Arehart, K.H.(1994). Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *Journal of the Acoustical Society of America*, 95(6), 3574-85.
- Oxenham, A.J.(2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends in amplification*, 12(4), 316–331. <https://doi.org/10.1177/1084713808325881>
- Bologna, W. J., Vaden, K. I., Ahlstrom, J. B., & Dubno, J. R. (2019). Age effects on the contributions of envelope and periodicity cues to recognition of interrupted speech in quiet and with a competing talker. *The Journal of the Acoustical Society of America*, 145(3), EL173-EL178.
- Stone, M. A., C Füllgrabe, & Moore, B..(2008). Benefit of high-rate envelope cues in vocoder processing: effect of number of channels and spectral region. *Journal of the Acoustical Society of America*, 124(4), 2272-82.
- Clinard, C.G., & Tremblay, K. L. (2013). Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *J. Am. Acad. Audiol.* 24, 590–599.
- Snyder, J. S., & Alain, C. (2005). Age-related changes in neural activity associated with concurrent vowel segregation. *Brain Res. Cogn. Brain Res.* 24, 492–499.
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87–90.
- Brown, C. A., & Bacon, S. P. (2010). Fundamental frequency and speech intelligibility in background noise. *Hearing research*, 266(1-2), 52–59.
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America*, 119(3), 1562-1573.
- Summerfield, Q., & Culling, J.F. (1992). Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency. *Journal of the Acoustical Society of America*, 92(4), 2317-2317.
- Qin, M. K., & Oxenham, A. J. (2006). Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *Journal of the Acoustical Society of America*, 119(4), 2417.
- Wingfield, A., Mccoy, S. L., Peelle, J. E., Tun, P. A., & Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *Journal of the American Academy of Audiology*, 17(7), 487-497.
- Assmann P., Summerfield Q. (2004) *The Perception of Speech Under Adverse Conditions*. In: *Speech Processing in the Auditory System*. Springer Handbook of Auditory Research, vol 18. Springer, New York, NY.

审稿人 2 意见:

意见 1: 本研究通过借助先进的语音处理技术和行为实验,揭示了 F0 曲线促进“鸡尾酒会场景”下语音感知/理解能力的老化特点。虽有研究结果虽然不能清楚地说明“老年人没有表现出 F0 促进效应”的具体原因,但是本研究为此问题的进一步探索做了很好的开端。我非常同意编委提出的建议: 对声学处理前后的语音特点进行更详细、直观地描述,对“老年人没有表现出 F0 促进效应”可能存在的原因展开讨论,以方便读者更好的理解。通过读此文的修改版,发现作者也对以上问题进行了较好地回答。但是,对于语音材料的声学特点,我还有一个小的问题: 初始语音分别合并 3 种 F0 轮廓后,语音的不同共振峰(如 F2,F3 的频谱特点)的分布特征在“合并”有怎样的变化,这一变化在不同的 F0 条件下是否相同。

回应:

非常感谢审稿专家对本文研究内容所提出的宝贵建议。
在以下的回应环节中将分为两个部分进行答复: 首先回答专家提出的问题,接着阐述由专家的问题所启发的进一步思考和对下一步工作的启示。

(1)由于汉语是声调语言,日常交流中使用的汉语语音的 F0 轮廓线索通常同时包含词汇声调和句子语调(林焘&王理嘉,1992)。在汉语语言中, F0 轮廓是单音节水平上声调感知的主要线索(Xu&Zhou,2011),还与句子语调和音高的感知有关。不同于前人的许多工作,本研究只对语调信息进行操作和修改,而保留了每个单音节的声调信息。
具体请详见本文 2.3 子标题下合成 F0 轮廓语音模型方法的描述(P5,第 5 段): “首先利用层次化音高目标模型和层次化时长模型对目标语句的 F0 轮廓进行建模,获取韵律信息的层次化特征,并对其中音调和语调之间的关联进行量化模拟;接着,分离出音调和句子语调;然后,保留音调,将语调进行参数化操作后,再将其重新叠加、平滑合成为连续的语音。这种操作的目的是在改变句子 F0 轮廓动态变化的同时保持其他韵律线索不变。”

由于在声学材料语音合成过程中,我们只对句子的语调进行扁平化或拉伸的调节操作,同时保留每个单音节的声调信息,尽可能地保持其他的韵律线索不变。因此,语音的不同共振峰的分布特征在“合并”后会尽量保持原有的共振峰模式不变。正常发音,基频不影响共振峰(跟口腔、鼻腔等相关)。如果是拉伸或压缩整段音频,采样点都变了,则会影响所有频率,包括共振峰。共振峰频率变化对频谱包络有影响,但与基频变化可以是互不相关的。

语音信号是调幅调频信号，频谱线索由信号的谐波结构、基频变化和共振峰共同构成。共振峰是识别元音内容的主要线索(吴宗济&林茂灿,1999)。共振峰的一个很重要的特性在于，它们在背景噪声中可以得到比较好的保持。例如，Assmann 和 Summerfield(2004)的研究表明，在言语识别中，外部的噪声会改变频谱的形状、改变其斜率、降低谱峰和谷之间的对比，但元音中共振峰的频率位置却可以较好地保持，能给听者提供不利条件下识别言语的线索。

因此，背景噪声下共振峰频率的动态模式可能会模糊听者感知 F0 轮廓动态变化所需的线索。那么老年听者无法感知动态追踪 F0 轮廓变化是否是因为对共振峰变化的干扰更敏感呢？

(2)非常感谢专家的意见，不仅启发了现有的讨论，也提供了下一步工作的潜在思路。

我们在讨论章节的第 10 页第 4 段增加了对老年人无法感知追踪动态 F0 轮廓线索有可能是因为对共振峰变化的干扰更敏感的讨论，具体增补内容请详见论文讨论部分(P10-P11)。增补修改后的内容如下所示：

“ 本文研究中使用的是汉语语句（相比于很多研究中所用的音节或元音-辅音-元音结构的刺激，无意义语句材料与正常的言语更为相似），言语信号在频谱结构和时间包络上快速不断地变化，这些复杂的模式很可能会干扰老年人对动态 F0 轮廓线索的感知和利用。因此，在不利条件下，错过这些线索的受益可能会增加老年人的言语识别难度。语音信号是调幅调频信号，频谱线索由信号的谐波结构、基频变化和共振峰共同构成。共振峰的一个重要特性在于它们在背景噪声中可以得到比较好的保持(Assmann & Summerfield, 2004)。共振峰频率的动态模式会伴随着信号的频谱和时间包络线而变化，这可能会模糊听者感知 F0 轮廓动态变化所需的线索(Souza et al.,2011)。所以，老年听者无法感知并追踪动态 F0 轮廓变化可能是因为对共振峰变化的干扰更敏感,从而对他们在言语识别中从动态 F0 线索中获益造成不利影响。因此，老龄化降低了对动态 F0 轮廓线索的感知是否源于对语音共振峰变化干扰的敏感性，值得未来研究的进一步关注。 ”

本部分内容中所涉及的相关参考文献如下所示：

- Assmann P., Summerfield Q. (2004) The Perception of Speech Under Adverse Conditions. In: *Speech Processing in the Auditory System. Springer Handbook of Auditory Research*, vol 18. Springer, New York, NY. https://doi.org/10.1007/0-387-21575-1_5
- Green T., Faulkner A., & Rosen S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. *The Journal of the Acoustical Society of America*, 112(5), 2155–2164.
- Assmann P. F. (1999). Fundamental frequency and the intelligibility of competing voices. In Ohala J. J., Hasegawa Y., Ohala M., Granville D., & Bailey A. C. (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 179–182). Oakland, CA: University of California Press.
- Shen, J., Wright, R., & Souza, P. E. (2016). On older listeners' ability to perceive dynamic pitch. *Journal of Speech Language & Hearing Research Jslhr*, 59(3), 572.
- Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research*, 235(1), 143–156.
- Souza, P., Arehart, K., Miller, C. W., & Muralimanohar, R. K. (2011). Effects of age on F0- discrimination and intonation perception in simulated electric and electro-acoustic hearing. *Ear and Hearing*, 32(1), 75–83
- Shen, J., & Souza, P. E. (2017). Do older listeners with hearing loss benefit from dynamic pitch for speech recognition in noise?. *American Journal of Audiology*, 26(3S), 462-466.
- 吴宗济&林茂灿，实验语音学概要，高等教育出版社，1989

第二轮

审稿人 1 意见：

意见 1：

实验设计为 2(年龄)×3(F0 线索类型)×4(信噪比)混合设计，但是仅考虑了 F0 轮廓和年龄这两个自变量，并未考虑信噪比对结果的影响。的线性混合模型分析中要同时考虑信噪比这一自变量。

回应：

感谢审稿专家的意见。

由于上一版修改稿的方法中我们直接计算了每名被试在每种实验条件下的言语识别阈限值用于统计分析。其中，识别阈限值指的是正确识别率达到 50% 时所对应的信噪比值，该值是当前评价言语识别能力的主要指标。因此言语识别阈限值也是

许多研究感兴趣的结果值，在很多研究(e.g. Buss et al.,2019; Shen & Souza, 2017; Shen & Souza,2019; Wu,2019)中都采用直接使用言语识别阈(SRT)值进行数据分析。

但根据专家的建议，鉴于实验设计为年龄×F0 线索类型×信噪比的三因素混合设计，仅考虑了 F0 轮廓和年龄这两个自变量，并未考虑信噪比对结果的影响，的确不够严谨。感谢审稿专家的建议，针对审稿专家的意见，我们对数据分析部分进行了相应的修改(详见修改后的论文中第三部分实验结果章节,P7-P9)，并将数据分析分为两个部分 1)对正确识别率的数据分析和 2)对识别阈限值的数据分析。

具体修改如下所示：

1)对正确识别率的数据分析(详见修改后的论文中第三部分实验结果章节, 3.1 对正确识别率的数据分析, P7-P8):

“实验使用 SPSS 统计软件对被试的正确识别百分率进行线性混合效应模型(linear mixed-effects models, LMMs; Baayen et al., 2008)统计分析。在本研究的模型中，包括 F0 轮廓类型(压扁、自然、拉伸)、年龄(年轻、老年)、信噪比水平(-4, 0, 4, 8 dB SPL)的固定效应，同时考虑被试随机效应。

全模型探讨了年龄、F0 轮廓类型、信噪比水平以及这些因素之间的相互作用对听者在言语掩蔽下的汉语言语识别成绩的影响。结果表明，年龄组($F(1,22)=194.150, p<0.001$)，F0 线索类型($F(2,242)=13.244, p<0.001$)和信噪比($F(3,242)=271.912, p<0.001$)的主效应都显著。显著的主效应表明，这三个因素都对汉语言语的可懂度存在贡献。在交互作用方面，F0 线索类型和信噪比两个因素的交互作用显著($F(6,242)=2.206, p=0.043$)；信噪比和年龄组两个因素的交互作用显著($F(3,242)=10.064, p<0.001$)；F0 线索类型和年龄组的交互作用也显著($F(2,242)=7.127, p=0.001$)，但三因素(F0 线索类型×信噪比×年龄)之间交互作用不显著($F(6,242)=1.269, p=0.272>0.1$)。

在各个年龄组下建立新模型，包括 F0 轮廓类型(压扁、自然、拉伸)、信噪比水平(-4, 0, 4, 8 dB SPL)的主效应和两重交互作用，对实验各因素的效应进行进一步的探测和分析。对年轻被试组的识别成绩统计分析结果表明，对于年轻被试组，F0 线索类型的主效应显著($F(2,121)=17.825, p<0.001$)，信噪比的主效应显著($F(3,121)=110.105, p<0.001$)。F0 线索类型和信噪比两个因素的交互作用显著($F(6,121)=2.696, p=0.017$)。对老年被试组的识别成绩统计分析结果表明：对老年被试组，信噪比的主效应显著($F(3,121)=178.271, p<0.001$)，但是 F0 线索类型的主效应($F(2,121)=0.962, p=0.385$)以及 F0 线索类型和信噪比之间的交互作用($F(6,121)=0.582, p=0.744$)均不显著。

在各个 F0 轮廓类型条件下建立新模型，包括年龄(年轻、老年)、信噪比水平(-4, 0, 4, 8 dB SPL)的主效应和两重交互作用，对实验各因素的效应做进一步的探测和分析。对于自然 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=152.071, p<0.001$)，信噪比类型的主效应($F(3,66)=73.034, p<0.001$)以及信噪比类型和年龄组之间的交互作用($F(3,66)=5.520, p=0.002$)都达到统计显著。对于扁平 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=117.759, p<0.001$)，信噪比类型的主效应($F(3,66)=113.783, p<0.001$)以及信噪比类型和年龄组之间的交互作用($F(3,66)=4.583, p=0.006$)均显著。对于拉伸 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=89.844, p<0.001$)，信噪比类型的主效应($F(3,66)=81.360, p<0.001$)显著，但是信噪比类型和年龄组之间的交互作用($F(3,66)=2.492, p=0.065$)不显著。对两个年龄组被试的识别成绩进一步分析表明，两组被试不管是在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下对目标语句的识别率都存在显著差异($ps<0.001$)。对年龄组的固定效应参数估计显示，在每种 F0 轮廓条件下老年被试组的言语识别成绩均显著低于年轻被试组(自然 F0 轮廓条件下： $\beta=0.267361, SE=0.056874, t=4.701, p<0.001$ ；扁平 F0 轮廓条件下： $\beta=0.170139, SE=0.053037, t=3.208, p=0.002$ ；拉伸 F0 轮廓条件下： $\beta=0.195602, SE=0.060321, t=3.243, p=0.002$)。

在各个信噪比水平条件下建立新模型，包括年龄(年轻、老年)、F0 轮廓类型(压扁、自然、拉伸)的主效应和两重交互作用，对实验各因素的效应做进一步的探测和分析。

对于信噪比为-4dB 水平条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=69.480, p<0.001$)，F0 类型的主效应显著($F(2,44)=7.103, p=0.002$)，以及 F0 类型和年龄组之间的交互作用($F(2,44)=4.382, p=0.018$)也显著。对于信噪比为 0dB 水平条件下的识别成绩统计分析结果表明：年龄组的主效应($F(1,22)=241.880, p<0.001$)，F0 类型的主效应 ($F(2,44)=8.903, p<0.001$)，以及 F0 类型和年龄组之间的交互作用 ($F(2,44)=4.508, p=0.015$)均显著。对于信噪比为 4dB 水平条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=53.887, p<0.001$)，F0 类型的主效应($F(2,44)=4.848, p=0.013$)显著，但 F0 类型和年龄组之间的交互作用($F(2,44)=0.834, p=0.441$)不显著。对于信噪比为 8dB 水平条件下的识别成绩统计分析结果表明：年龄组的主效应显著($F(1,22)=34.106, p<0.001$)，但 F0 类型的主效应($F(2,44)=0.795, p=0.458$)以及 F0 类型和年龄组之间的交互作用($F(2,44)=1.299, p=0.283$)都不显著。”

2)对识别阈限值的数据分析 (详见修改后的论文中第三部分实验结果章节, 3.2 对识别阈限值的数据分析, P8-P9):

“ 实验使用 SPSS 统计软件, 采用线性混合效应模型对被试的识别阈限值进行统计分析。在本研究的模型中, 包括 F0 轮廓类型条件(压扁、自然、拉伸)、年龄组(年轻、老年)的固定效应, 同时考虑被试随机效应。

实验分析结果显示, F0 轮廓类型($F(2,44)=7.145, p=0.002$)和年龄组组间差异($F(1,22)=73.214, p < 0.001$)对被试的识别成绩均有显著影响, F0 线索类型和年龄组的交互作用显著($F(2,44)=4.931, p=0.012$)。固定效应参数估计进一步显示, 老年被试组的言语识别成绩明显差于年轻被试组($\beta=-7.731E0, SE=0.903553, t=-8.557, p < 0.001$)。

来自年轻被试组识别阈限值的数据分析显示, 在对 F0 线索操作的影响方面, 不同 F0 轮廓类型对被试识别成绩的影响显著($F(2,22)= 6.874, p=0.005$)。在模型中将自然 F0 轮廓条件作为基线条件, 对 F0 轮廓被调节的条件和自然 F0 轮廓条件之间的识别成绩进行比较。结果显示, 与自然动态 F0 轮廓线索条件相比, 压扁 F0 轮廓条件下的言语识别表现明显较差($\beta=2.9158E0, SE=0.876849, t=3.325, p=0.003$); 拉伸 F0 轮廓条件下的言语识别成绩也比自然动态 F0 轮廓条件下的显著较低($\beta=2.7033E0, SE=0.876849, t=3.083, p=0.005$)。但压扁 F0 轮廓与拉伸 F0 条件下的言语识别成绩之间差异不显著($\beta=0.212500, SE=0.876849, t=0.242, p=0.811$)。

对老年被试组识别阈限值的数据分析表明, 不同 F0 轮廓线索类型对老年被试的言语识别成绩的影响不显著, $F(2,22)= 0.498, p=0.614$ 。压扁 F0 轮廓条件($\beta=0.339760, SE=0.340560, t=0.998, p=0.329$)或拉伸 F0 轮廓条件($\beta=0.163173, SE=0.340560, t=0.479, p=0.637$)下的言语识别成绩与自然动态 F0 轮廓线索条件下的言语识别成绩之间差异均不显著。”

本部分内容中所涉及的参考文献如下所示:

- Wu, M. H. (2019). Effect of F0 contour on perception of Mandarin Chinese speech against masking. *Plos One*, 14(1).
- Shen, J., & Souza, P. E. (2017). Do older listeners with hearing loss benefit from dynamic pitch for speech recognition in noise?. *American Journal of Audiology*, 26(3S), 462-466.
- Buss, E., Hodge, S. E., Calandruccio, L., Leibold, L. J., & Grose, J. H. (2019). Masked Sentence Recognition in Children, Young Adults, and Older Adults: Age-Dependent Effects of Semantic Context and Masker Type. *Ear and hearing*, 40(5), 1117-1126.
- Shen, J., & Souza, P. E. (2019). The ability to glimpse dynamic pitch in noise by younger and older listeners. *Journal of the Acoustical Society of America*, 146(3), E1232-E1237.

意见 2:

研究中测量了青年组和老年组的纯听阈限, 建议将纯听阈限作为协变量纳入数据分析, 排除纯听阈限的不同对结果可能产生的影响。

回应:

感谢审稿专家的意见。

文章中的第二部分实验方法中 2.1 被试章节展示了两个年龄组被试在各个频率下的纯音测听阈限。该部分测量纯听阈限是为了筛选被试, 即要求参加实验的参与者必须是双耳平衡且对称的, 不存在传导性损伤。

在许多研究中, 用纯音平均值测量外周听力被作为个体听力的一种测量方法(e.g.

陈艾婷, 2012), 但也值得注意的是, 这种测量方法可能无法捕捉听者感知动态 F0 线索的能力(Shen & Souza, 2017)。具体来说, 老年人感知动态 F0 线索的能力可能会因各种阈上听力缺陷而下降, 这其中包括时间编码不良(Grose & Mamo, 2010; Hopkins & Moore, 2011)、频率选择性降低(Hopkins & Moore, 2011), 以及频率调制的退化神经表征(Clinard & Cotter, 2015)。因此, 不应完全排除老龄相关性外周听觉能力下降在损害利用 F0 轮廓线索促进嘈杂环境下言语感知中的作用。假若要完全排除可能产生的影响, 除了 audiogram hearing thresholds 外, 可能还需要测量包括 ABR; FP; GIN; DPOAE 等等的指标。这将是另外一个开放话题, 值得未来进一步的研究。

非常感谢审稿专家建议的启发, 我们也在文章的第四部分讨论章节中的最后一段未来进展中补充探讨了相关问题(详见修改后的论文中第四部分讨论章节, P12):

“值得注意的是, 在前人的研究中, 老年人年龄相关性听力下降与声音辨别能力(Helfer & Freyman, 2008)或声音精细结构保持的初始听觉“记忆”能力(Huang et al., 2008)没有显著相关性; 老年人在利用知觉空间分离线索从信息掩蔽中释放目标语音的能力并不比年轻人差 (Helfer & Freyman, 2008; Li, Daneman, Qi, & Schneider, 2004)。尽管如此, 不应完全排除年龄相关性听力损失在损害利用 F0 轮廓线索以促进噪声下言语感知方面的作用。在未来的工作中, 研究阈上听觉能力测量和动态 F0 线索获益之间的关系具有重要性。从临床角度来看, 这项工作还可以扩展到包括一组患有听力损失的年轻人或老年人, 以研究听力损失对噪声下言语识别的动态 F0 线索获益程度的影响。”

本部分内容涉及的参考文献如下所示:

Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research*, 323, 91-98.

Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G.(2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37), 15516-15521.

Grose, J.H. , & Mamo,S.K..(2010). Processing of temporal fine structure as a function of age. *Ear Hear*, 31(6), 755-760.

Hopkins, K., & Moore, B.. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, 130(1), 334.

肖容, 梁丹丹, 李善鹏. (2020). 汉语普通话声调感知的老年化效应: 来自 ERP 的证据. *心理学报*, 52(1), 1-11.

陈艾婷. (2012). 噪声下言语测听材料的开发和相关研究——噪声下普通话言语测听材料和语句识别算法的研究. (Doctoral dissertation, 第四军医大学).

意见 3:

样本量是否足够, 作者通过 G power 进行了计算, 建议将计算结果作为脚注增加到被试介绍部分。

回应:

感谢审稿专家的建议。已经根据专家的建议将 G*power 软件的计算结果作为脚注增加到被试介绍部分(详见修改后的论文中第二部分方法章节,P8-P11)。具体的脚注增补内容如下:

“采用 G*Power 3.1.9 软件对研究所需样本量进行先验(priori)分析。以中等效应量 $f=0.3$, 统计检验力 0.95 ($\alpha=0.05$) 估计最小样本量为每组 10 个样本(Total sample size:20), 表明本文中当前的被试量(青年组和老年组,每组 12 个)是足够的。”

意见 4:

有关无意义句子的构成, 需要报告每个句子中两个名词和一个动词出现的概率标准, 较低的可预测性的具体数值是多少。

回应:

感谢审稿专家的意见。已经根据专家的建议, 我们在文章第二部分方法章节的 2.3 刺激材料, a.中文“无意义”语句章节, 对每个句子中两个名词和一个动词出现的概率标准进行报告。具体修改如下:

“为了确保这些句子都是无意义的, 实验对每个句子中所用的两个名词和动词在所选的《人民日报》数据库(1994-2002)中在一个正常句子中同时出现的概率都做了严格限制。

只有当无意义语句中的三个关键词在所选数据库同一个正常句子中共现的概率为零时，该语句才会被用作实验材料(详见 Yang et al., 2007)。”

意见 5:

因为结果分析中所存在的问题，缺乏对三个自变量及其交互作用的讨论。需要根据数据分析结果重新撰写讨论部分，讨论 F0 线索类型、信噪比和年龄这些变量及其交互作用对结果产生的影响。

回应:

感谢审稿专家的意见。已经根据专家的建议重新写了讨论章节的内容，并在讨论部分增补并讨论 F0 线索类型、信噪比和年龄这些变量及其交互作用对结果产生的影响(详见修改后的论文中第四部分讨论章节,P10-P11)，具体修改内容如下所示:

“实验研究评估了老年被试组与年轻被试组在言语掩蔽下聆听具有自然 F0 轮廓的汉语语句与拉伸的或扁平的 F0 轮廓的相应语句的言语识别成绩。数据分析结果发现 F0 轮廓类型、信噪比水平和年龄因素都会对听者言语识别成绩产生影响。研究结果显示信噪比对年轻被试和老年被试都存在显著的影响($p < 0.001$)。不管是老年被试组还是年轻被试组，在言语掩蔽下关键词的识别率都随着信噪比的提高而升高。具体地，在本实验的所有条件中，关键词的识别率都随着信噪比从-4 dB 增加到 8 dB 而呈单调增加的趋势，没有表现出任何平台效应。言语掩蔽的干扰声音下，这种目标识别率的单调递增趋势是与前人所报告的结果一致(Brungart et al., 2001; Li et al., 2004; Wu, 2019)。

研究结果还显示，对于年轻被试组，不管是在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下，年轻被试对目标语句的识别率都随信噪比的降低而降低。数据分析结果表明在每种 F0 轮廓条件下，信噪比对目标识别率的作用都显著($p < 0.001$)。对于老年被试组，实验结果也显示不管是在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下，老年被试对目标语句的识别率都随信噪比的变化而变化。数据分析表明信噪比对识别率的作用不管是在自然 F0 轮廓条件($p < 0.001$)下还是在被调节的 F0 轮廓线索的条件($p < 0.001$)下都显著。这个结果与年轻被试的结果一致。然而，与年轻被试的结果不同的是，对老年被试而言，自然动态 F0 轮廓条件产生的去掩蔽作用与扁平的 F0 轮廓条件或拉伸的 F0 轮廓条件产生的去掩蔽作用相当($p > 0.05$)。而年轻被试在自然动态 F0 轮廓条件产生的去掩蔽作用与扁平的 F0 轮廓条件($p < 0.001$) 或拉伸的 F0 轮廓条件($p < 0.001$)产生的去掩蔽作用存在显著差异。

具体实验测量结果显示，言语掩蔽下给年轻被试呈现中文无意义语句时，具有自然动态 F0 轮廓的语句比具有扁平的 F0 轮廓或拉伸 F0 轮廓的语句都表现出较好的言语识别成绩。说明具有自然 F0 轮廓的条件比具有拉伸或扁平 F0 轮廓的条件更好地被感知。这一结论发现证明了自然 F0 轮廓线索对年轻听者在噪声下言语识别中的重要作用。研究结果也表明了对 F0 轮廓线索的操作会对听者在言语掩蔽下的汉语言语识别表现产生有害影响。本研究的结果与英语语音感知研究的文献(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)中的结论一致，表明与自然的 F0 轮廓模式相比，被调节操作的动态 F0 线索对噪声中的言语感知有负面影响。同时，本研究结果也支持了 Patel 等人(2010)和 Wu(2019)中关于 F0 轮廓的自然动态变化对背景噪声中的汉语言语识别具有重要贡献的论述。

对老年被试的数据分析结果显示，老年被试在言语掩蔽下聆听具有自然 F0 轮廓的汉语语句与拉伸的或扁平的 F0 轮廓的相应句子呈现出同样低的言语识别成绩。研究结果表明，扁平或拉伸 F0 轮廓的条件没有能帮助老年人在言语掩蔽环境下抵抗信息掩蔽，自然动态变化的 F0 轮廓的条件也没有比对 F0 轮廓操作后的条件更能抵抗信息掩蔽，说明在多人说话的嘈杂环境下老年听者可能难以利用动态 F0 轮廓线索以促进对掩

蔽下的目标言语识别。”

意见 6:

关于本文的讨论，作者基于老年人的阴性结果做了大量的讨论，当基于阴性结果做出解释时要谨慎。是否可以通过一些统计指标说明阴性结果是可靠的？

回应:

感谢审稿专家的意见。

上一轮修改稿对“老年人没有表现出 FO 促进效应”可能存在的原因展开大量讨论，探讨了多人语音掩蔽下老年人很可能对 FO 线索的探测和跟踪出现了困难。鉴于专家的意见，我们意识到“基于老年人的阴性结果做了大量的讨论”这种解释方式的不足，因此，我们缩减了该领域讨论的篇幅，并补充探讨了对潜在的有益方式的探索（详见修改后的论文中第四部分讨论章节,P11-P12）。

意见 7:

关于“瞥见”效应的讨论与本研究的关系需要更进一步的阐述，很难理解“瞥见”效应与本研究所操纵的自变量之间有什么关系。

回应:

感谢专家的意见，我们重新思考了关于“瞥见”效应的相应讨论，认为在这里增补探讨“基频 FO 轮廓线索对促进老年人目标言语识别去掩蔽作用变得有限可能还和老年人“瞥见”能力下降有关”并不合适。因此，我们对上一轮修改稿中第 10 页第三段关于“瞥见”效应的讨论进行了整段删除。

意见 8:

讨论中用到“脆弱”，应该如何理解？表现在哪些方面？什么机制导致的“脆弱”？本研究的哪些研究结果支持了上述观点。

回应:

感谢审稿专家的意见。

根据审稿专家的前几点建议，我们已经重写了讨论部分，原文中有关脆弱描述的相应句子已经被改写。由于文章阐述要面向读者，尽量通俗易懂，为了避免讨论部分晦涩难懂，新修改后的讨论章节中没有用到“脆弱”的修辞。

第三轮

审稿人 1 意见:

感谢作者认真仔细的修改。再次建议作者增加一些统计指标比如贝叶斯值说明阴性结果的可靠性。

回应:

感谢审稿专家对论文修改的认可，非常感谢审稿专家对本文研究内容所提出的有益建议。

根据专家的具体建议，我们已经增加了贝叶斯值统计指标来说明阴性结果的可靠性。并将分析结果附在论文中 3 实验结果部分末尾，具体修改过程如下所示（详情请参考修改后的论文中第三部分实验结果中的 3.2 章节，P9-P10，第 5、6、7、8 段）:

贝叶斯值的统计指标量化了备择假设 H1 和零假设 H0 的相对证据，亦即比较了在 H1 成立的前提下观察到观测数据的概率和在 H0 成立的前提下观察到数据的概率。

在实际研究中能够对零假设 H_0 和备择假设 H_1 提供量化的证据具有重要性,它可以直观让研究者区分出有证据表明没有效应和没有证据表明有效应这两种情况。

正文中补充修改的具体内容如下:

“

为了进一步探究数据支持两个主效应以及主效应间交互作用显著的可靠性,使用统计软件 JASP (<https://jasp-stats.org/>, JASP Team 2019) (Wagenmakers, Love, Marsman, et al., 2018)对数据进行贝叶斯分析检验,分析产生的贝叶斯因子(BF_{10})描述了数据对假设的支持程度,能够量化备择假设 H_1 和零假设 H_0 的相对证据。

实验使用 JASP 的默认分布,对两个年龄被试组的识别阈限值数据进行贝叶斯重复测量方差分析,结果显示年龄组主效应的 BF_{10} 值是 635658.207, 根据贝叶斯因子的决策标准(Jeffreys, 1961), 该结果表明有极强的证据支持 H_1 ; F0 轮廓类型主效应的 BF_{10} 值是 8.655, 有中等强度的证据支持 H_1 ; 年龄组和 F0 轮廓类型的主效应之和的 BF_{10} 值是 $5.589e+6$, 表明与单独的 F0 轮廓主效应相比,当加入年龄的效应时,数据对 H_1 的支持提高了大约 $6.46e+05$ 倍;在此基础上,当把年龄组和 F0 轮廓类型的交互作用也增加到模型中时, BF_{10} 值是 $2.254e+7$, 表明有极强的证据支持 H_1 。

对年轻被试组的数据进行贝叶斯重复测量方差分析,结果显示贝叶斯因子 $BF_{10}=12.324$; 说明有较强的证据支持 H_1 , 即年轻人 F0 轮廓类型的主效应显著。对老年被试组的数据进行贝叶斯重复测量方差分析,结果显示贝叶斯因子 $BF_{10}=0.265$; 说明存在中等强度的证据支持 H_0 , 即老年人的 F0 轮廓线索类型的主效应不显著。

进一步的简单效应检验使用 JASP 默认分布对老年被试在不同 F0 轮廓类型条件下的数据两两进行贝叶斯配对样本 t 检验,分析得出的贝叶斯因子数值(压扁的 F0 轮廓条件和自然 F0 轮廓条件, BF_{10} 为 0.418; 自然 F0 轮廓条件和拉伸 F0 轮廓条件, BF_{10} 为 0.339; 压扁 F0 轮廓条件和拉伸 F0 轮廓条件, BF_{10} 为 0.315), 根据 Jeffreys (1961) 提出的分类标准,表明当前的数据有较弱的证据支持了 H_0 , 也说明观察到的老年被试的数据相对于备择假设 H_1 更加支持没有效应的假设 H_0 , 即老年被试的识别成绩在不同 F0 轮廓类型的条件之间没有显著差异。 ”

本部分内容相关的参考文献如下所示:

- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.
- Wetzels, R., & Wagenmakers, E.J. (2012). A default bayesian hypothesis test for correlations and partial correlations. *Psychon Bull Rev*, 19(6), 1057-1064.
- Wagenmakers, E.J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25(1), 58-76.
- JASP Team (2019). JASP (Version 0.11.1) [Computer software].
- Marsman, M., & Wagenmakers, E.J. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology*, 14, 545-555.
- Wagenmakers, E.J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, 25, 169-176.
- 胡传鹏, 孔祥祯, & 彭凯平. (2018). 贝叶斯因子及其在 JASP 中的实现. *心理科学进展*, 26(6), 951-965

编委意见: 同意发表。

主编意见: 建议接受