心 理 学 报
第1期 ACTA PSYCHOLOGICA SINICA 1993年

发音人和语境变异与声调知觉[1)]

杨玉芳 梁晓光
中国科学院心理研究所，北京，100012

摘 要

本文用 Garner 快速分类法研究了发音人规范化和语境补偿过程与声调知觉的关系。得到的主要结果是：(1)发音人变化对声调识别产生了显著的 Garner 干扰；声调变化对男女声判别产生了同样的干扰。(2) 语境变化对声调识别也产生了显著的 Garner 干扰。按照 Garner 方法的假设，这说明发音人规范化和语境变异补偿与声调识别之间是相互联系和相互影响的。(3)虽然发音人和语境同时变化引起的声调识别反应时延长约等于它们单独变化时反应时延长量之和，但正确率下降远大于单独变化时下降量之和。说明发音人和语境变异对声调识别的干扰不是独立的，它们之间存在相互作用。这种作用增加了声调识别的难度。

关键词：发音人规范化、语境变异补偿，Garner 快速分类法，声调知觉

一、引 言

在话语中，不同发音人所发以及同一发音人在不同语境下所发的同一语音，无论在波形上还是语音分析参数上都存在很大差异。这些差异在话语中往往代表有用的信息，使得在噪声背景下快速的言语交流成为可能。但对语音识别来说，它们又成为干扰源。

在言语交流中，听者面临的问题是，一方面要利用这里变异携带的语音学和语言学信息，另一方面要补偿由于发音人和语境引起的变异。言语知觉研究的一个方面是研究这种补偿过程的确切机理以及它与其它言语加工水平间的关系。Blandon、Henton 和 Pickering(1984)[1]提出了元音规范化听觉模型，把基频作为确认发音人的主要线索，用以建立与发音人有关的元音知觉空间。Miller 等 (1989)[2] 研究了讲话速度对语音范畴知觉结构的影响，认为听者在适应语速时涉及到声学信号与语音结构整体之间的重新匹配。还有一些人研究了发音人规范化与其它言语加工水平之间的关系。Mullennix 等 (1990)[3] 证明，补偿由发音人引起的变异和语音知觉是两个性质不同而又紧密相关的过程，对言语信号中任何一种信息的选择性注意，都会影响到对另一种信息的加工。但这种影响是不对称的，发音人变化对辅音识别的干扰远大于辅音变化对发音人识别的干扰。他们把这种关系称为有条件的平行关系。Jonhson(1990)[4]研究了元音对比和规范化的关系，发现元音的确认函数受到环境因素的影响，环境因素的作用增强了元音规范化效应。环境影

* 本研究得到国家自然科学基金资助。
1) 本文于 1992 年 6 月 4 日收到。

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

响发生在确认讲话者的水平上，而不是元音音色的水平上。元音规范化要参照所确认的发音人。

上述研究工作对于了解发音人规范化和语境补偿过程的机理及其与其它加工水平之间的关系是十分有益的。但同时可以看出，其中使用的语音材料是十分有限的。可以想象，发音人规范化与语音识别的关系会因音位性质不同而不同，不能一概而论。因为辅音、元音和声调这三种音位与作为发音人规范化主要线索的基频之间的关系是不同的。语境补偿作为另一种预处理过程，有与规范化过程不同的目的任务，与音位识别之间可能存在相似或不相似的关系。本文所要研究的问题是：(1)声调知觉与发音人规范化和语境补偿过程之间的关系是什么？(2)当听者必须同时对两种变异进行补偿时，两种补偿过程间有无相互作用？相互作用的结果是什么？

采用 Garner 快速分类法[5]研究这些问题。这种方法的原理是，比较听者在两个变量的三种组合条件下的操作，看其能否有选择地注意某一变量而忽略另一变量的干扰，以此推断对两种变量的加工是有联系的还是相互独立的。三种组合条件是：(1)控制条件——非注意变量保持恒定，注意变量变化；(2)正交条件——注意和非注意变量同时变化；(3)相关条件——注意和非注意变量的某些水平成对匹配。与控制条件相比，正交条件下操作水平的下降，包括反应时的延长和正确率的下降，称为 Garner 干扰；在相关条件下，由于区别因素的增加而产生的操作水平的提高，称为 Garner 增益。

二、实 验 和 结 果

声调主要取决于基频随时间变化的模式，同其他音位一样，在话语中会随发音人和语境产生变化。变化可能表现在音域、时长和基频曲线的形状等方面。在本实验中考虑的语境变异只是二字词连读时产生的声调变化。语境变化共有八种，在一、二、三、四声前用 C_1，C_2，C_3 和 C_4 表示，在一、二、三、四声后用 C_5、C_6，C_7 和 C_8 表示。考虑切分时的方便，采用音节 da 负载声调。由两名男发音人和两名女发音人将两个 da 音节四声的十六种组合以二字词的形式读出。经过 4kHz 低通滤波、放大后，以 10kHz 采样。然后用数字编辑

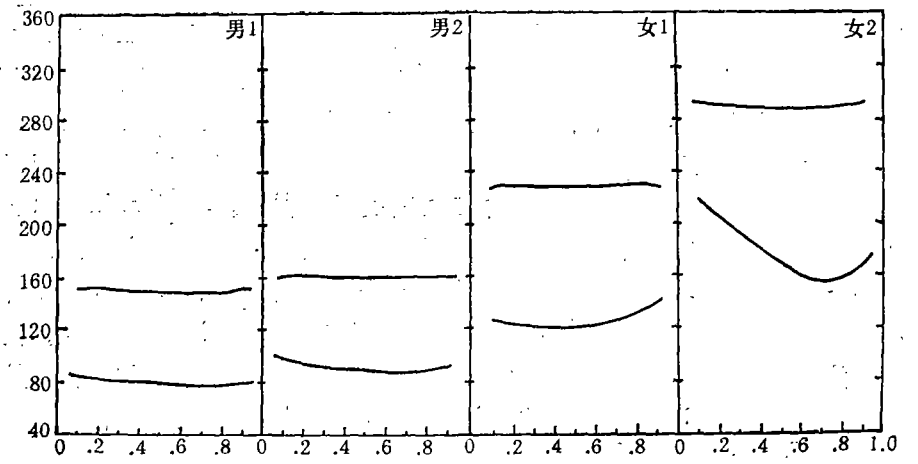

图1 四名发音人在 C_1 下的一声和 C_2 下的三声的基频曲线(横坐标为归一化的时长)

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

方法把两个音节切开，共得到 128 个音节。为说明四名发音人音域的差别和语境的影响，图 1 中给出了四人在 C_1 下的一声和 C_2 下的三声的基频曲线；图 2 中给出了一名男发音人在八种语境下所发的四声的基频曲线。这些曲线都是用线性预测分析得到的。

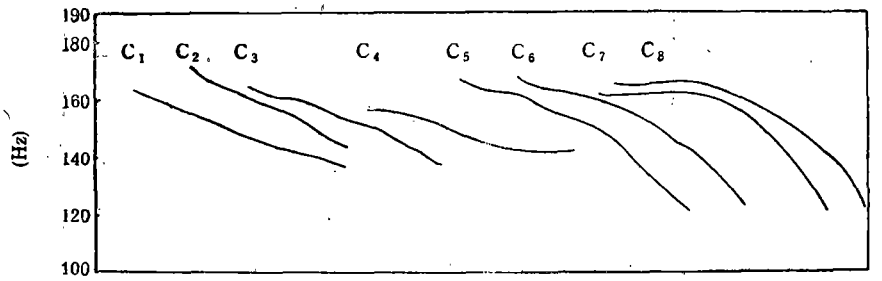

图 2　一名男发音人在八种语境下发的第四声的基频曲线(横坐标为时间，单位长度为 10ms)

实验分成三部分。实验一研究发音人规范化和声调识别间的关系；实验二研究语境变化对声调识别的影响；实验三探讨同时有发音人和语境变化条件下的声调识别。参加实验的共有十八名听者，分为三组。第一组的顺序是实验一、二、三；第二组为实验二、一、三；第三组为实验三、二、一。听者都是青年学生，年龄在 18—24 岁间，讲普通话，无听力问题和耳疾。

全部实验由一个框架程序控制，将指定的刺激文件以随机顺序调出，作 D/A 转换后，通过滤波和放大，由 DL-1 耳机传送给听者。刺激间隔为 1 秒。听者通过键盘作反应，程序同时记下反应和反应时。完成全部实验约需 2 小时。实验在隔声室中进行，背景噪声不大于 40dB。

1. 实验一：发音人变异和声调知觉

实验一研究发音人规范化过程与声调知觉的关系，使用的语音材料是四名发音人在语境 C_2，即二声 dá 作为二字词后音节时所发的一、二、三、四声，共 16 个音节。实验分成两部分。

第一部分用 Garner 快速分类法研究发音人规范化过程与声调知觉的关系。听者的任务是识别声调，用数字键 1、2、3、4 作反应。有三种实验条件：(1)在控制条件下，一个发音人所发的四个声调各重复八次，三十二个音节随机回放，形成一节。共有四节，每节间隔为 1 秒。(2)在正交条件下，四个发音人发的四个声调混合，重复八次，随机回放，构成一节。(3)在相关条件下，第一名男发音人的一声，第二名男发音人的二声，第一名女发音人的三声和第二名女发音人的四声各重复三十二次，混合后随机回放，共 128 个音节。三种条件的顺序在被试组间做到平衡。

第二部分用同样方法研究声调变异对发音人性别判断的影响。听者的任务是区分男声和女声，若为男声，按“M”键；若为女声，按“F”键。在控制条件下，四名发音人所发的一种声调重复八次，随机回放，组成一节。共有四节，每节 32 个音节。在正交条件下，四人所发的四种声调重复八次，随机回放，组成一节，共 128 个音节。相关条件同第一部分的相关条件。

两部分实验中在三种条件下听者的反应时和正确率由表 1 和表 2 给出。

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表1 识别声调和男女声的反应时 单位：ms

识别任务		条件			干扰	增益
		控制	正交	相关		
声调	M SD	428.7 62.9	465.2 88.4	404.7 59.7	36.5	24.0
男女声	M SD	297.9 36.6	336.9 57.7	317.0 67.1	39.0	-19.1

表2 识别声调和男女声的正确率（%）

识别任务		条件			干扰	增益
		控制	正交	相关		
声调	M SD	96.7 2.3	93.3 3.8	97.3 2.8	3.4	0.6
男女声	M SD	95.9 2.8	93.5 3.7	92.0 5.4	2.4	-3.9

从表1可以看出，识别声调的实验结果是，在正交条件下反应时长于控制条件，相关条件下反应时短于控制条件。识别男女声的反应时，正交条件长于控制条件，相关条件也长于控制条件。发音人变化和声调变化所产生的Garner干扰分别为36.5ms和33.8ms。二维ANOVA分析结果是，识别声调和男女声的反应时差异显著($F=88.143, P=0.000$)，三种实验条件下反应时差异也显著($F=4.462, P=0.014$)，但识别任务和实验条件间交互作用不显著($F=1.282, P=0.282$)。pos hoc检验结果是，识别声调时，三种条件间反应时差异均显著($P<0.014$)。识别男女声时，正交条件与控制条件，相关条件与控制条件的反应时差异均显著($P<0.002$)，但正交条件与相关条件间差异不显著($P=0.098$)。

以正确率为指标得到的结果与以反应时为指标的结果一致。操作任务和实验条件的作用均显著($P<0.006$)，且二者间的交互作用也显著($P=0.017$)。

上述实验结果说明，发音人变异对声调识别产生了Garner干扰，声调变化对男女声识别也产生了Garner干扰。但在相关条件下发音人变异产生了明显的Garner增益，而声调变化对男女声识别无显著影响。

2. 实验二：语境变异和声调知觉

实验二用Garner快速分类法研究语境补偿过程与声调知觉的关系，使用的语音材料是一名男发音人所发的全部32个音节。控制条件有八节，每一节对应一种语境下所发的四个声调，每个声调重复八次。正交条件只有一节，使用八个语境下的四个声调，每个声调重复八次，共256个音节。在相关条件下，使用语境C_1、C_5下的一声，C_2，C_6下的二声，C_3、C_7下的三声和C_4、C_8下的四声。每个声调重复32次。

实验中听者的任务是识别声调，用数字键1、2、3、4作反应。控制程序同时记录反应和反应时。表3和表4分别给出有语境变异时三种条件下声调识别的反应时和正确率。

ANOVA分析结果是，三种实验条件间反应时差异不十分显著($F=2.226, P=0.057$)，

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 3　语境变化时声调识别的反应时（ms）

反应时	条件			干扰	增益
	控制	正交	相关		
M	427.7	474.3	444.3	46.6	-16.6
SD	52.6	91.7	75.1		

表 4　语境变化时声调识别的正确率（%）

正确率	条件			干扰	增益
	控制	正交	相关		
M	94.6	91.8	92.7	2.8	-1.9
SD	2.8	3.1	5.1		

正确率差异也不显著($F=2.524, P=0.090$)。但 pos hoc 检验结果说明，控制和正交两种条件间差异显著(对于反应时，$F=4.112, P=0.047$；对于正确率，$F=4.797, P=0.033$)，相关条件与另外两种条件间差异不显著($P\geqslant0.511$)。

实验数据说明，语境变化增加了声调识别的反应时，降低了正确率。根据 Garner 方法的假设，这说明在本实验条件下，语境补偿过程与声调识别是相互联系，相互影响的。

3. 实验三：发音人和语境同时变化

该实验测量在发音人和语境同时变化时声调识别的反应时和正确率。通过与没有非注意变量(控制条件)和只有一个非注意变量(实验一、二的正交条件)时的操作水平作对比，通过两个变量对声调识别的影响及交互作用的分析，研究发音人规范化和语境变异补偿这两个过程间的关系。

实验使用了四名发音人在八种语境下所发的全部 128 个音节，每个音节重复两次，随机回放。要求听者识别声调。结果是，声调识别的平均反应时为 498.8ms，标准差为51.2ms；正确率为 83.2%，标准差为 5.2%。

表 5 给出了实验三的结果。为便于分析比较，同时列出了实验一、二的控制条件和正交条件下声调识别反应时和正确率。其中第一行数据为两个控制条件下结果的平均值。

表 5　不同非注意变量条件下声调识别的反应时、正确率和干扰

非注意变量	反应时（ms）		正确率(%)	
	反应时	干扰	正确率	干扰
无	428.0	——	95.7	——
发音人	465.2	37.2	93.3	2.4
语境	474.3	46.3	92.7	3.0
发音人和语境	498.8	70.8	83.2	12.5

对不同非注意变量条件下声调识别反应时和正确率作方差分析。结果是，发音人和语境同时变化时声调识别的反应时和正确率都与两种控制条件下的差异显著($P<0.001$)。发音人变化、语境变化、发音人和语境同时变化这三种条件下的反应时差异不显著($F=0.495, P=0.613$)，但正确率差异极为显著($F=34.783, P=0.000$)。pos hoc 检验结果是，发

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

音人和语境单独变化这两种条件间差异不显著($F=0.124$, $P=0.726$)，而这两种条件与两个非注意变量同时变化时差异显著(P 均为0.000)。

对上述三种条件下的 Garner 干扰进行方差分析。结果是，发音人变化和语境变化所产生的 Garner 干扰，无论是以反应时作指标还是以正确率作指标，差异均不显著。它们与语境和发音人同时变化所产生的干扰差异极为显著(P 均为 0.000)。

当以反应时作指标时，发音人和语境同时变化所产生的干扰约等于它们单独变化时产生的干扰之和；以正确率作指标时，却大大超过了单独产生的干扰之和。这说明，对声调识别来说，发音人和语境同时变化大大增加了识别难度。这不仅是由于要同时处理和补偿两种不同的变异源，还可能由于两种变量和两种过程间的相互作用。

三、讨 论

本实验用 Garner 快速分类法研究了发音人规范化和语境补偿过程与声调知觉的关系。得到的主要结果是：(1)发音人变化对声调识别产生了显著的 Garner 干扰；声调变化也对男女声判别产生了同样程度的干扰。按照 Garner 方法的假设，这说明发音人规范化和声调识别这两个加工过程是相互联系，相互影响的，而且影响是对称的。(2)语境变化对声调识别也产生了显著的 Garner 干扰，语境变异补偿与声调识别之间也存在相互联系。(3)发音人和语境同时变化引起的反应时延长约等于两个因素单独变化引起的反应时延长之和，正确率下降远大于单独变化时下降量之和。可以说，发音人和语境变异间存在交互作用，交互作用的结果是增加了声调识别的难度。

Mullennix 等(1990) 曾专门研究过发音人规范化与辅音识别的关系。得到的结果之一是，发音人变化对辅音识别的干扰远大于辅音变化对男女声判断的干扰，即二者之间的影响是非对称的。这与本实验观察到的发音人规范化与声调识别间关系不同。这可能是由于二者使用了不同的识别任务。男女声区别在一定程度上可以说在于基频；声调区别主要在于基频曲线的高低和升降，二者间存在有机联系。因此声调变化对男女声判别的干扰远大于辅音变化。这也说明不同加工过程之间的联系方式和紧密程度是有差别的。

语境变化本身十分复杂。变异来源、补偿机制及其与语音识别的关系可能都与发音人规范化过程不同。在语流中，相邻音节的声调通过时长、音域、曲线形状变化和对比来传达语调和节奏信息。听者在识别声调时要参照句子的这些整体特征。本实验仅研究了二字词内部连读引起的声调变化。这种特定的语境变化对声调识别产生了显著的Garner干扰。按照 Garner 方法的假设，这说明语境补偿与声调识别间是有联系的。但与发音人规范化不同的是，在相关条件下发音人变化增加了对比因素，有助于声调识别，产生了 Garner 增益；语境变化则对声调识别没有显著影响。有的研究者认为，相关条件下有无 Garner 增益，指示了两个过程间关系是串行还是并行。这是否符合本实验情况，有待进一步考察。由于语音材料的限制，实验中没有探讨声调变化对语境判断的影响，这留待以后的研究加以解决。

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

参　考　文　献

[1] Blandon, A.,Hendon, C., and Pickering, J. B., Towards an auditory theory of speaker normalizition, Language and Communication, 1984, 4, 59—69.
[2] Miller, J., Interactions in processing segmental and suprasegmental features of speech, Perception and Psychophysics, 1989, 24, 175—180.
[3] Mullennix, J. W., and Pisoni, D. B., Stimulus variability and processing dependences in speech perception, Perception and Psychophisics, 1990, 47(4), 379—390.
[4] Jonhson, K., Contrast and normalization in vowel perception, Journal of Phonetics, 1990, 18, 229—254.
[5] Garner, W.R., The processing of information and structure, Potomac, MD: Erlbaum, 1974.

THE EFFECTS OF THE SPEAKER AND CONTEXTUAL VARIATION ON TONE PERCEPTION

Yang Yufang　　Liang Xiaoguang
Institute of Psychology, Chinese Academy of Sciences

Abstract

Three experiments were carried out to investigate the effects of normalization of the speaker and contextual compensation on tone perception, using the paradigm of Garner speed classification. It was shown that the speaker and contextual variation produced significant Garner Interference (GI) with tone perception, and tone variation produced the same degree of GI with voice discrimination, when the speaker and context variated in an experiment bolck, GI indexed by reaction time was approximately equal to the sum of GI produced separately by the speaker and contextual variation, GI indexed by percentage of correct responses was greater than the sum of GI by the two variations.

Key words: Speaker normalization, contextual compensation, the paradigm of Garner speed classification, tone perception.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net