第23卷 第5期 物 理 学 报 Vol. 23, No. 5
1974 年 9 月 ACTA PHYSICA SINICA Sept., 1974

音节清晰度与音位清晰度之间的统计关系*

张 家 骕
(中国科学院物理研究所)

提 要

在大量普通话清晰度试验的基础上，根据汉语音节结构的具体分析，认为语言的内部信息——结构规律在言语的主观识别过程中是十分重要的因素．这特别表现在言语信号受到较大的干扰或失真，言语信号的物理特性已不能提供充分的识别条件时，内部信息发生较大的作用．

本文提出了考虑到这一内部信息作用的音节清晰度与音位清晰度之间的统计关系．这一关系，比 Fletcher 与 Steinberg 所建立的统计关系，更好地符合大量试验的结果．

言语清晰度试验和言语可懂度理论，不仅是评价言语传递系统并指导其提高工作质量的主要手段，而且它已成为研究听觉器官的一种心理物理方法．

早在 1908 年 Rayleigh[1] 研究电话机时，便注意到了言语清晰度问题．他为使电话机能分出 f 和 s 而进行了不少试验．随后，1910 年，Campbell[2] 作了专门的清晰度试验．到 1929 年，Fletcher 与 Steinberg[3] 则对清晰度试验进行了比较系统而广泛的研究．同时，提出了语音清晰度与音节清晰度之间的统计关系，至今人们还广为利用．

我们在数种传递条件下(改变接收声压级，在白噪声掩蔽下接收，言语经过高、低通滤波系统以及几种条件的混合)，进行了大量清晰度试验，处理了总数约一百万个语言单位(包括音节、单词、单句)．最后导出了汉语的清晰度指数，测试了不同传递条件对可懂度的影响，并对各种语言单位的清晰度试验得分之间的关系进行了统计分析．

基于对语言特性(物理的与社会的)的认识，以及对普通话语音结构的具体分析，本文提出了音节清晰度与音位清晰度之间的新的统计关系．

分析结果表明，新关系更好地符合大量试验结果．

一、问题的提出

几乎和发展清晰度试验的同时，人们就企图建立各种语言单位的清晰度试验得分之间的统计关系，因为进行清晰度试验是相当繁重的工作．此外，分析各种语言单位的清晰度试验得分之间的关系，也为揭示人的听觉在接收言语——这一人类所特有的信号系统时的内在规律，提供一定的线索．

Fletcher 与 Steinberg 所建立的音节清晰度与音位清晰度之间的统计关系为

* 1973 年 4 月 16 日收到．

$$S = A_1L + A_2L^2 + \cdots\cdots + A_nL^n, \tag{1}$$

其中 S——音节清晰度；L——语音清晰度；A_n——由 n 个语音组成的音节在音节总数中所占的比例．在这一关系中，包含着这样的假定，即各语音的正确接收是纯机会性的，并且在接收过程中各语音是相互独立的．

不同语言的试验结果[4-6]，按照这一公式进行计算时，都存在实测值与计算值在清晰度较低时有较大的偏离(见图 1)．特别值得注意的是，这一偏差具有一恒定的偏向，即实测值常大于计算值．Fletcher[4] 早就发现了这一现象，但没有解决这一问题，只是把它称之为"X"影响．在处理汉语清晰度试验结果时，也发现了类似现象．甚至由日本人用英语进行的清晰度试验[7]也出现了同样现象．这是为什么呢？

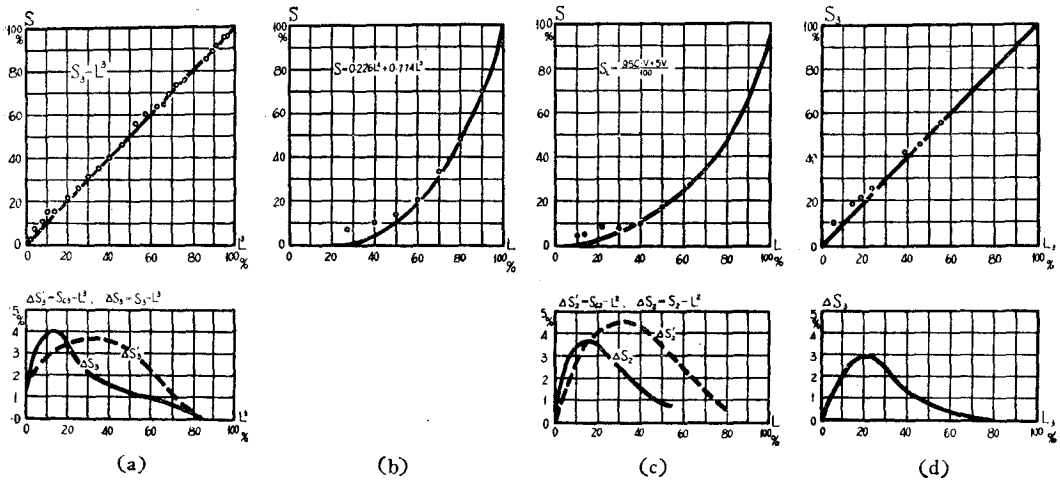

图 1 几种语言的音节清晰度 S 和 ΔS 与语音清晰度 L 的关系

(a) 引自 Fletcher (英语)；
(b) 引自 Bikov (俄语)；
(c) 引自 Miura (日语)；
(d) 引自 Torii (英语)

二、数 学 模 型

Fletcher 与 Steinberg 公式的假定条件是不符合实际情况的．不论从发音器官发音，还是听觉器官在知觉过程中，各语音之间都不是相互无关的．因而听音人在接收言语信号时，各语音的正确接收并不是独立事件．

在语言结构中，各种语言单位之间并不是随机组合的，而是存在着一定的语音学和语法上的规则的．这种规则在人们长期运用中，作为一种语言的内部信息而储存于记忆当中．听音人在收到言语信号的物理特性以后，还加上已储存于记忆当中的社会特性作出其响应．

对于汉语，这种规则在音节结构当中便明显地表现了出来．汉语音节包含声、韵、调三种成分．它们之间的组合是有比较严格的规律的．我们在编制语音平衡的汉语清晰度试验表(KX 表)时，以及分析处理试验记录时，均根据"汉语拼音方案"规定的声、韵、调进行的．由于汉语的声调，在所试验的各种常见的传递条件下，是不发生错误的(清晰度近

于 100%)，因此，我们主要分析音节清晰度与声母、韵母清晰度的关系.

在普通话中，声母、韵母的结合关系见表 1.

表 1　普通话音节结构分析表

声母＼韵母	开口呼	齐齿呼	合口呼	撮口呼
O*	有	有	有	有
b p m	有	有	u	无
f	有	无	u	无
d t	有	有	有	无
n l	有	有	有	有
g k h	有	无	有	无
j q x	无	有	无	有
zh ch sh r	有	无	有	无
z c s	有	无	有	无

注：1. O* 代表零声母.
　　2. 合口呼的 u 表示只与 u 相拼不与 u 领头的复韵母相拼.

从表 1 中可见，汉语音节的组合不是随机的，因而它被听音人接收时，其中各成分也不是独立的.

当我们把清晰度得分看作是正确接收语言单位的概率时，可以采用概率理论中的 Bayes 公式：

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{P(B)} \tag{2}$$

和

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i), \tag{2a}$$

作为处理这一问题的数学模型.

假设，听音人在作出一个音节响应这一事件 B 时，可以在 A_1, A_2, A_3, $\cdots A_n$ 个条件下进行，那么，$P(B)$ 便是在各种条件下作出一个音节响应的全概率.

听音人在作出一个汉语音节响应时，不外下列三种情况(音位清晰度为 L 时)：

1. 声母、韵母全听对了，其概率为 $P(A_1)=L^2$;
2. 声母、韵母只听对其中之一，其概率为 $P(A_2)=2L(1-L)$;
3. 声母、韵母全听错了，其概率为 $P(A_3)=(1-L)^2$.

这样，$P(B|A_i)$ 便是在上述某一情况下，写出一个音节的概率. 当取 $P(B|A_i)$ 是声母、韵母全听对了，写出一个音节的概率时，则 $P(B|A_i)=P(B|A_1)=1$，那么 $P(A_i|B)=P(A_1|B)$ 便是听音人已经作出了一个音节响应，而这个音节的声韵母都对了，即作出了一个正确的响应的事后概率，这也就是音节清晰度 S 了.

不管听音人所收到的语音受到了何等的畸变或干扰，他所作出的响应只能是一个普通话音节，这是音节表所预先规定的. 所以，虽然他听错了一个声母或韵母，还是有可能写出一个普通话音节的. 设此时写出一个音节的概率 $P(B|A_2)=q_{12}$. 显然，q_{12} 与普通话音节结构有关. 它可用普通话中现有音节数在声韵母的可能组合中所占的比例来表示.

在声韵母全听错的情况下，写出一个普通话音节的概率为 $P(B|A_3) = Lq_{12}$.

我们把 q_{12} 叫作音节结构常数. 根据统计计算，$q_{12} = 0.87$.

这样，我们便可以将式(2)写成

$$S = \frac{L^2}{L^2 + q_{12} \cdot 2L(1-L) + Lq_{12} \cdot (1-L)^2}, \tag{3}$$

其中 L——音位清晰度$\left(L = \frac{I+F}{2}\right)$; I——声母清晰度; F——韵母清晰度; S——音节清晰度.

三、结果及讨论

在几种不同传递条件下，包括改变接收声压级，改变信-噪比 S/N，经过高、低通滤波系统以及几种情况的混合，进行了清晰度试验. 所得的音节清晰度和音位清晰度的实测值(圆点)与计算值(实线)的比较示于图 2.

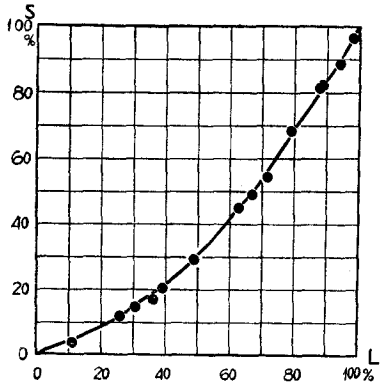

图 2 汉语音节清晰度与音位清晰度的关系

从图 2 中可以看到，计算值与实测值符合得是比较好的，并消除了图 1 中其它作者的恒定偏差现象.

现在我们尝试用新的关系式去校正日语和英语的清晰度试验结果与计算值的关系.

1. 对于日语来说，音节是 C-V 结构，可采用式(3)来处理. 根据日语音节的具体情况，可假定式(3)中的 $q_{12} = 1$，式(3)则成为

$$S_{c2} = \frac{L^2}{L^2 + 2L(1-L) + L(1-L)^2} = \frac{L^2}{1-(1-L)^3}. \tag{4}$$

可以看到，S_{c2} 恒大于 L^2，当 L 较小时更为显著.

今取新旧计算关系的差值:

$$\Delta S_{c2} = S_{c2} - L^2 = L^2\left[\frac{1}{1-(1-L)^3} - 1\right]. \tag{5}$$

现在把 ΔS_{c2} 与 L 的关系也绘于图 1(c) 上(虚线)，可以看到两条曲线是颇为近似的.

2. 对于英语，按 Fletcher 与 Steinberg 所用的 C-V-C 音节表，可将式(2)写成

$$S_{c3}=\frac{L^3}{L^3+q_{23}\cdot 3L^2(1-L)+q_{13}\cdot 3L(1-L)^2+Lq_{13}\cdot(1-L)^3},\tag{6}$$

其中　S_{c3}——按新关系三音音节清晰度的计算值；L——语音清晰度 $\left(L=\frac{2c+V}{3}\right)$；$q_{23}$——已正确接收两个语音写出一个三音音节的概率；$q_{13}$——已正确接收一个语音写出一个三音音节的概率；$Lq_{13}$——三个语音均接收错误写出一个三音音节的概率.

首先，为简化分析起见，令 $q_{23}=1$，$q_{13}=1$，这样式(6)可写成

$$S_{c3}=\frac{L^3}{1-(1-L)^4},\tag{7}$$

可见，即使是语音任意组成 C—V—C 都构成一个音节时，也恒存在 $S_{c3}>L^3$ 的关系.

其次，为了进一步作些定量估计，现在考查 Fletcher 与 Steinberg 的编表方法. 他们共选取了 25 个辅音，11 个元音；其中三个辅音 w, h, y 只用作 C—V—C 音节的前辅音，另三个辅音 zh, ng, st 只用作后辅音，其余均前后适用，因而共有 22 个前辅音和 22 个后辅音. 在不计入他们舍去的一些俚语等音节的情况下，只考虑上述约束组成音节时，计算得到

$$q_{13}=0.77,\ q_{23}=0.92.$$

将 q_{13}，q_{23} 代入式(6)，就可算出新公式与旧公式的差值

$$\Delta S_{c3}=S_{c3}-L^3.\tag{8}$$

现将 ΔS_{c3} 也绘于图 1(a)（虚线）上. 可以看到，由新公式给出的校正曲线与实测值和旧公式计算值之差是颇为近似的，它们的数值大小和趋势都是一致的. 这就从另一方面证明了我们所建立的新的统计关系的可靠性和普遍性.

由此可以看出，语言结构规律直接影响着听音人的主观识别. 我们在将这种规律用于分析汉语单词可懂度与音节清晰度的关系时，也得到了较好的结果.

四、结　　论

根据所得结果，可得出如下结论：

1. 不同语言单位的清晰度试验得分之间存在着互相依赖的统计关系；
2. 它们之间的关系决定于语言系统的结构规律——内部信息；
3. 当言语信号的物理特性受到严重损害，清晰度很低时，听音人便以所收到的部分信号为核心，依照语言结构规律而向外扩展，以完成一个响应，这时内部信息起着重要作用；
4. 我们给出的新关系，还只能是对言语知觉这一复杂过程中的一个环节的近似描述.

清晰度试验工作是集体的大量的劳动结果，这一试验工作是在中国科学院物理研究所九室许多同志参加下共同完成的.

参 考 文 献

[1] Lord, Rayleigh, *Phil. Mag., 6th Series,* **16** (1908), 235.
[2] G. A. Campbell, *Phil. Mag., 6th Series,* **19** (1910), 152.
[3] H. Fletcher and J. C. Steinberg, *J.A.S.A.,* **1** (1929), 1.
[4] H. Fletcher, Speech and Hearing in Communication (1953 New York).
[5] Yu. S. Bikov, Intelligibility Theory and Increase the Efficiency of Radio Communication, (1959, Moscow & Leningrad) (in Russian).
[6] T. Miura, Quality of Speech Communication (1955) (in Japanese).
[7] N. Torii, *J.A.S.J.,* **11** (1955), 108.

ON THE STATISTICAL RELATION BETWEEN THE SYLLABLE ARTICULATION AND THE PHONEME ARTICULATION

ZHANG JIA-LU

(Institute of Physics, Academia Sinica)

ABSTRACT

The internal information, i.e. the syllable formation rule, of the language plays an important role in the statistical relation between the syllable articulation and the phoneme articulation, especially when the articulation scores are low. This is taken account in deriving an expression for the relation between the syllable articulation and the phoneme articulation scores, which agrees well with the experimental results of standard colloquial Chinese and accounts for the differences between theory and experiment for English and Japanese by previous authors.