

听觉对视觉时间知觉的影响及其机制

翟梦蝶

苏州大学教育学院，苏州

摘要 | 人们对不同通道时间信息的感知存在差异，听觉通道的时间知觉更加精确且稳定，因而同步呈现视觉与听觉时间信息时，视觉刺激的主观时间知觉会被任务无关的听觉刺激吸引，主要表现为：对视觉刺激呈现时刻的感知向同步听觉刺激靠拢；视觉刺激的主观时距会受同步听觉刺激的时距影响而拉长或缩短。此外，视听双通道刺激的时间判断精度高于视觉刺激单独呈现的情形。这种听觉对视觉时间知觉的影响可能是由于听觉刺激加快了视觉时间的脉冲发送速率，也可能基于模式开关开放与闭合潜伏期的改变。

关键词 | 时间知觉；时间腹语术；视听整合

Copyright © 2021 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

中世纪哲学家奥古斯丁在《忏悔录》中言及：“什么是时间？如果没有人问我，我知道。如果我要向发问者解释，我则一无所知。”古往今来，物理学家们执着于“测量”时间，心理学家则侧重人们对时间的主观感知。生活中交谈、运动到并欣赏音乐、睡眠等方方面面都离不开对时间的正确感知。时间知觉包括时序知觉与时距知觉，“时序知觉”即对事件发生先后顺序的知觉；“时距知觉”则指事件的持续时间^[1]。时距又分为“空时距”与“实时距”，“空时距”指一个事件与另一个事件之间空白的时距；而“实时距”指代某一事件持续存在的时距^[2]，无论是空时距还是实时距，人们的主观时距知觉会受到诸多因素的影响。先后有众多研究者探讨了记忆^[3]、个体差异^[4]、运动^[5]、注意^[6]、情绪^[7]、刺激呈现的通道^[8]等因素对人们主观时距的影响。其中，时距信息呈现的通道一直是时间知觉领域研究的焦点。由于大脑中迄今为止并未发现专门的时间信息感受器，大脑处理时间信息是基于一个中枢的机制还是多个分散的机制至今尚未成定论，而研究不同通道来源的时距信息有助于鉴别大脑的时间信息加工机制。

作者简介：翟梦蝶，苏州大学教育学院在读硕士研究生。

文章引用：翟梦蝶. 听觉对视觉时间知觉的影响及其机制[J]. 中国心理学前沿, 2021, 3(5): 427-432.

<https://doi.org/10.35534/pc.0305052>

2 视觉与听觉时间知觉的差异

刺激呈现通道的差异影响人们对时距主观上的感知 (perceived duration) 也影响时距判断的精度 (precision) [9]。听觉刺激在时间判断中一般居于主导地位 [10]，听觉时距信息的判断精度高于等长的视觉时距，且对于客观上相等的视觉与听觉时距，听觉刺激的主观时距往往长于视觉刺激 [8]。此外，听觉刺激的时间判断过程也比视觉刺激稳定，视觉时间信息的判断极易受到其他因素的影响 (如运动)，而听觉时间特性的判断则十分稳定 [5]。Schütz 和 Morrone (2010) 发现 [11]，在被试进行平滑眼球追踪运动 (Smooth pursuit eye movement) 时，视觉空时距的主观知觉出现压缩，而听觉空时距的主观时距与其客观时距没有显著差异。进一步说明了听觉信息在时间判断中的主导地位。

由于视觉与听觉刺激在时间判断中的差异，当一段空时距由视觉与听觉信息编码时，也会导致主观时距的差异。具体来说，当一段间隔时间的“开始 (onset)”由听觉刺激编码而“结束 (offset)”由视觉刺激编码，其主观时距会长于由视觉标记时间间隔开始、由听觉刺激标记时间间隔结束的条件 [12]。并且，该现象也可以扩展到触觉领域，即由听觉刺激标记开始的空时距长于由触觉刺激标记的空时距 [10]。

时距知觉领域中一个重要的理论是“标量期望理论 (Scalar Expectancy Theory, SET)”，即假定大脑中存在一个“内部时钟 (internal clock)”，其中包含起搏器 (pacemaker)、模式开关 (mode switch)、累加器 (accumulator) 三个部分 [8]。当一段时距计时开始，模式开关闭合，起搏器开始发送脉冲 (pulse)，计时结束后开关再次打开，脉冲进入累加器，并与记忆中的参考时距进行对比，进而推算出主观上的时距 [2]。而听觉刺激之所以拥有更高的时间分辨率，可能是由于听觉刺激的计时过程中，起搏器发送脉冲的速率更快 [13]；但 Chen 和 Yeh (2009) 认为 [8]，模式开关闭合与开放的潜伏期差异也有可能导致在利用听觉刺激进行时间判断时，敏感性与稳定性跟高。Kanai, Lloyd, Buetti 和 Walsh (2011) 通过经颅磁刺激 (Transcranial Magnetic Stimulation, TMS) 刺激初级视觉皮层与初级听觉皮层提出 [14]，损伤初级视觉皮层只会影响被试对视觉时间信息的判断，而损伤初级听觉皮层可以同时干扰听觉视觉时间信息的感知。进而 Kanai 等 (2011) 认为 [14]，由于听觉刺激具有更高的时间分辨率，人们在对视觉时间信息编码时，会将视觉时间信息转为听觉编码，从而进行时间判断，换言之，在进行时间判断时，听觉通道是时间信息加工的“默认通道 (default modality)” [9]。

3 听觉对视觉时间知觉的影响

想象你正观看一场精彩绝伦的电影。屏幕上的罗密欧与朱丽叶生离死别，伴随着伤感婉转的配乐，让人沉迷。虽然大脑加工视觉与听觉信息的速度存在差异，但观看电影的你却并未“意识到”到这细微的差异，这是由于大脑将不同来源的信息进行了“跨通道整合 (Cross-modal integration)” [15]。生活中我们经常需要同时加工几种不同通道来源的信息，若这些信息在时空特性上存在差异，大脑在加工这些信息时会尽力做到“最佳整合 (optimal integration)”，使不同来源的信息整合为一个独立连贯的事件 [16]。虽然这种跨通道的整合有利于我们正确而快速地感知事件，但也会导致我们对某一特定通道的信息知觉产生扭曲。

在空间领域，最为人所知的是“腹语术 (ventriloquism)”现象。在腹语术表演中，演讲者唇部不动而手里的布偶嘴巴一张一合，观众会错误地认为那些让人捧腹的言语出自这只布偶口中，这是由于人们在进行空间定位时，由于视觉刺激具有较高的空间敏感性，听觉刺激的空间位置会被视觉刺激的空间位

置“吸引”，进而人们产生定位偏差^[17]。与之相反，在时间领域，听觉刺激的时间分辨率高于视觉刺激，因此听觉刺激的时间特性会对视觉刺激的主观时间知觉产生影响，该现象称为“时间腹语术(Temporal ventriloquism)”现象^[18]。

听觉刺激会影响人们对视觉刺激呈现时刻的主观感知以及对时距的主观感知。在闪光滞后效应(Flash-lag effect)中，当闪光与一个运动物体在某个位置共同出现时，这个闪光知觉上会落后于运动物体，而当这个闪光伴随着一个声音共同呈现，闪光滞后效应显著降低^[18]，说明单个听觉刺激对运动视觉刺激在特定时刻呈现时间的感知(perceived onset)有吸引作用。此外，若一个声音先于闪光出现，另一个声音晚于闪光出现(即两个声音居于两个视觉刺激外侧)，视觉刺激的主观间隔会被两个无关的听觉刺激拉的“更远”，因而对两个视觉刺激的时序感知更为精确^[19]。与此同时，若视觉时距呈现的同时伴随呈现一个听觉时距，人们对该视觉时距的主观感知长于它单独呈现的情形^[20]。值得注意的是，即使客观上该听觉刺激与视觉刺激相等，视听双通道呈现时，该双通道刺激的主观时距长于单通道视觉刺激的主观时距，且这种效应在实时距中更为突出^[21]，换言之，听觉等长的空时距并不一定会影响视觉空时距的主观感知^[20]。

上文提及，听觉刺激与视觉刺激不仅在主观呈现时刻、主观时距的感知方面存在差异，二者时间判断的精度也存在差异。然而，同步呈现的听觉刺激并不一定会导致视觉刺激时间判断精度上升。Romei, De Haas, Mok 和 Driver (2011)采用时间比较法^[22]，给被试呈现两个连续的视觉刺激，并且，视觉刺激呈现时可能同时呈现长于、短于或等于视觉时距的听觉时距，结果发现，匹配的听觉刺激会提升视觉刺激的时间判断精度。Hartcher-O'Brien 和 Alais (2011)让被试听两段 1250 ms 的声音刺激^[23]，其中任何一个都有可能出现目标刺激(视觉或视听刺激)，被试须判断哪段声音刺激中包含目标刺激。结果发现，同步呈现的声音刺激并未提升视觉目标的判断精度，因此 Hartcher-O'Brien 和 Alais (2011)认为^[23]，视听时间特性的整合只能达到“次最佳整合(Suboptimal integration)”，即虽然声音的呈现提高了视觉刺激的探测率，但并未提升视觉刺激的时间分辨率。但 Romei 等(2011)^[22]与 Hartcher-O'Brien 和 Alais (2011)^[23]实验结果的差异可能源于实验刺激的差异，Romei 等(2011)采用的是视觉与听觉实时距^[22]，而 Hartcher-O'Brien 和 Alais (2011)本质是探讨单个刺激(即呈现时间点)的探测^[23]，单个短暂听觉刺激的唤醒度可能远远低于一段听觉实时距^[24]。Hartcher-O'Brien 等(2014)则同样采用了视觉实时距以及时间比较法^[16]，并在 Romei 等(2011)^[22]实验的基础上，改变同步听觉刺激的信噪比并建立数学模型，证明听觉刺激可以提升视觉刺激的时间辨别力，具体表现为，视听双通道刺激的精度高于视觉单通道刺激的时间精度且低于听觉单通道刺激的时间精度，符合贝叶斯“最大似然估计(Maximum Likelihood Estimation, MLE)”模型，即视听刺激的时间特性整合遵循“最佳整合”模型，与其余特性(如空间特性)的跨通道整合过程一致。

4 听觉调节视觉时间知觉的机制

在时序判断中，听觉对视觉时序判断的影响可能由于视听整合导致了两个视觉刺激主观上向听觉刺激的呈现时间靠拢，进而二者相差更远，使得两个视觉刺激的最小可觉差(Just Noticeable Difference, JND)降低^[25]。但在时距判断中，需要注意的是，如果需要判断一段时距的长短(无论空时距还是实时距)，

只有当这段时距呈现完毕后, 被试才能判断出刚刚呈现的时距信息“多长”, 因此, 时距判断总是带有回溯 (post hoc) 的性质, 因而在探讨视听时距信息的整合时, 需要注意这两个信息的整合并非实时进行, 而是在所有信息呈现完毕后才能整合。对此, Hartcher-O'Brien 等 (2014) 认为^[16], 视听时距信息的整合可能有两种机制: 1、大脑首先表征一段视觉时距与一段听觉时距, 而后将两种不同的时距加以整合, 即“时距整合 (interval integration)” ; 2、大脑首先将表征时距开始的视觉与听觉标记 (marker) 整合, 而后将表征时距结束的视觉与听觉标记 (marker) 整合, 最后对这段由双通道刺激标记的时距进行表征, 即标记整合 (marker integration)。由于第二种整合方式包含了两个整合过程而第一种整合方式只包含一个整合过程, 因此从整合的结果来看, 时距整合的变化性 (variance) 更小, 精度更高。

这两种假设的机制反应在标量期望理论中, 其一是听觉可能加速了视觉时间信息编码时发放脉冲的速率; 其二是听觉刺激缩短了视觉模式开关的潜伏期, 使得开关开放与闭合的时间差异增加。听觉加速视觉刺激的脉冲速率的确可能解释部分时距膨胀现象, 但 Klink 等 (2011) 研究发现, 当听觉刺激客观上短于视觉刺激的时距时, 视觉刺激的主观时距会出现压缩, 若听觉仅仅基于加快视觉时间的脉冲发送速率进而影响视觉刺激的主观时距, 那么显然不能解释为何较短的听觉时距可以压缩视觉刺激的主观时距。

此外在实证研究方面, 虽然关于实时距的研究很好地支持了时距整合的观点, 但空时距的实证研究表明, 视听时距整合时也会涉及部分标记整合的过程。Morein-Zamir 等 (2003) 的实验二中^[25], 其中一个听觉刺激与视觉刺激客观上同时呈现 (即客观上对齐), 而另一个听觉刺激, 可能领先第一个视觉刺激也可能落后于第二个视觉刺激。结果发现, 当第一个听觉刺激领先视觉刺激而第二个听觉刺激与第二个视觉刺激同时呈现时, 听觉刺激并没有提升视觉刺激的时间分辨率; 而当第一个听觉刺激与视觉刺激同时呈现, 第二个听觉刺激晚于视觉刺激呈现时, 视觉刺激的时间分辨率上升, 只是上升的幅度较小。Shi, Chen 和 Müller (2010) 通过探究单个声音对“视觉似动错觉现象 (Ternus effect)” 的影响表明^[26], 当第二个声音落后于第二帧视觉刺激 30 ms 时, 被试回答“整体 (group) 移动” 的比例更高 (即听觉影响了视觉刺激的主观感知)。这两个实验均表明, 单个声音的呈现也可以吸引对视觉刺激呈现时刻的感知, 进而将两个视觉刺激的时间间隔主观上拉长。而若视听整合的过程完全基于“时距整合”, 单个听觉刺激理应不会影响视觉刺激的主观时间间隔。而在动物实验中得出了相反的结论。Feenders 和 Klump (2018) 让欧椋鸟 (starling) 判断哪一侧 LED 灯先亮并选择相应的栖木^[27], 研究结果发现, 当单个声音领先时, 欧椋鸟选择正确的比例上升; 而当单个声音落后于视觉刺激时, 欧椋鸟的行为表现没有显著差异。Feenders 和 Klump (2018) 认为^[27], 单个声音领先时行为上产生的微弱差异是由于率先出现的声音的“警示 (alerting)” 作用, 而第二个声音并未提升视觉时间判断精度这一现象 (与 Morein-Zamir 等人研究不符), 是由于单个声音与两个闪光违背了“统一体假说 (unity assumption)”, 即欧椋鸟并没有将这个单独的声音与其中任何一个闪光当作是同一来源的刺激, 进而 Feenders 和 Klump (2018) 认为^[27], 只有当被试可以判定听觉与视觉刺激是“统一体” 时, 才可以发生视听整合。虽然该实验的研究结果与 Morein-Zamir 等 (2003)^[25]、Shi 等 (2010)^[26] 的研究不完全一致, 但依旧认同视听时间特性的整合过程会发生一定程度的时间标记整合, 只是被试须主观上认同这两个时间标记时间上足够接近进而认为这两个刺激本质上是同一来源的刺激。

虽然一系列研究认为视觉与听觉时间特性的整合过程中包含了时距整合与时间标记整合两个过程, 但从行为结果来看, 这两个过程并不能完全分离, 大脑在整合视听时间信息的过程中相对依赖于时距信

息,但也权衡了时间标记的信息,未来研究可尝试通过改变两种信息的信噪比进一步研究二者如何互相权衡,进而探究人们如何不依赖外部计时器却可以相对精确地加工时间信息的机制。

参考文献

- [1] 凤四海, 黄希庭. 时间知觉理论和实验范型 [J]. 心理科学, 2004 (05): 1157-1160.
- [2] Wearden J H, Edwards H, Fakhri M, et al. Why “Sounds Are Judged Longer Than Lights”: Application of a Model of the Internal Clock in Humans [J]. *The Quarterly Journal of Experimental Psychology*, 1998, 51 (2): 97-120.
- [3] Ogden R S, Wearden J H, Jones L A. Are memories for duration modality specific? [J]. *Quarterly Journal of Experimental Psychology*, 2010, 63 (1): 65-80.
- [4] Corcoran A W, Groot C, Bruno A, et al. Individual differences in first-and second-order temporal judgment [J]. *Plos one*, 2018, 13 (2): e0191422.
- [5] Merchant H, Yarrow K. How the motor system both encodes and influences our sense of time [J]. *Current Opinion in Behavioral Sciences*, 2016, 8: 22-27.
- [6] Spence C, Squire S. Multisensory integration: maintaining the perception of synchrony [J]. *Current Biology*, 2003, 13 (13): R519-R521.
- [7] Bocanegra B R, Zeelenberg R. Emotion-induced trade-offs in spatiotemporal vision [J]. *Journal of Experimental Psychology: General*, 2011, 140 (2): 272.
- [8] Chen K M, Yeh S L. Asymmetric cross-modal effects in time perception [J]. *Acta psychologica*, 2009, 130 (3): 225-234.
- [9] Retsa C, Naish P, Bekinschtein T, et al. Temporal judgments in multi-sensory space [J]. *Neuropsychologia*, 2016, 8: 101-112.
- [10] Grondin S. Why studying intermodal duration discrimination matters [J]. *Frontiers in psychology*, 2014, 5: 628.
- [11] Schütz A C, Morrone M C. Compression of time during smooth pursuit eye movements [J]. *Vision research*, 2010, 50 (24): 2702-2713.
- [12] Mayer K M, Di Luca M, Ernst M O. Duration perception in crossmodally-defined intervals [J]. *Acta psychologica*, 2014, 147: 2-9.
- [13] Klink P C, Montijn J S, van Wezel R J A. Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates [J]. *Attention, Perception, & Psychophysics*, 2011, 73 (1): 219-236.
- [14] Kanai R, Lloyd H, Buetti D, et al. Modality-independent role of the primary auditory cortex in time estimation [J]. *Experimental Brain Research*, 2011, 209 (3): 465-471.
- [15] Keetels M, Stekelenburg J, Vroomen J. Auditory grouping occurs prior to intersensory pairing: evidence from temporal ventriloquism [J]. *Experimental Brain Research*, 2007, 180 (3): 449.
- [16] Hartcher-O'Brien J, Di Luca M, Ernst M O. The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion [J]. *Plos one*, 2014, 9 (3): e89339.
- [17] Bischoff M, Walter B, Blecker C R, et al. Utilizing the ventriloquism-effect to investigate audio-visual binding [J]. *Neuropsychologia*, 2007, 45 (3): 578-586.

- [18] Vroomen J, de Gelder B. Temporal ventriloquism: sound modulates the flash-lag effect [J] . Journal of Experimental Psychology: Human Perception and Performance, 2004, 30 (3) : 513.
- [19] Shimojo S, Scheier C, Nijhawan R, et al. Beyond perceptual modality: Auditory effects on visual perception [J] . Acoustical Science and Technology, 2001, 22 (2) : 61-67.
- [20] De la Rosa M D, Bausenhardt K . Multimodal integration of interval duration: Temporal ventriloquism or changes in pacemaker rate? [J] . Timing & Time Perception, 2013, 1 (2) : 189-215.
- [21] Wearden J H, Norton R, Martin S, et al. Internal clock processes and the filled-duration illusion [J] . Journal of Experimental Psychology: Human Perception and Performance, 2007, 33 (3) : 716.
- [22] Romei V, De Haas B, Mok R M, et al. Auditory stimulus timing influences perceived duration of co-occurring visual stimuli [J] . Frontiers in psychology, 2011, 2: 215.
- [23] Hartcher-O'Brien J, Alais D. Temporal ventriloquism in a purely temporal context [J] . Journal of Experimental Psychology: Human Perception and Performance, 2011, 37 (5) : 1383.
- [24] Lake J I, LaBar K S, Meck W H. Emotional modulation of interval timing and time perception [J] . Neuroscience & Biobehavioral Reviews, 2016, 64: 403-420.
- [25] Morein-Zamir S, Soto-Faraco S, Kingstone A. Auditory capture of vision: examining temporal ventriloquism [J] . Cognitive Brain Research, 2003, 17 (1) : 154-163.
- [26] Shi Z, Chen L, Müller H J. Auditory temporal modulation of the visual Ternus effect: the influence of time interval [J] . Experimental Brain Research, 2010, 203 (4) : 723-735.
- [27] Feenders G, Klump G M. Violation of the unity assumption disrupts temporal ventriloquism effect in starlings [J] . Frontiers in psychology, 2018, 9: 1386.

Influences of Auditory Stimulus on Visual Time Perception and Its Mechanism

Zhai Mengdie

School of Education, Soochow University, Suzhou

Abstract: The sensory modality influences the subjective duration of the stimulus. Time information presented in auditory modality appears more accurate and stable. Interestingly, the auditory stimulus would capture visually perceived onset, and the visual subjective duration will be prolonged or compressed by synchronous auditory duration. Moreover, the precision of the cross-modal stimulus is higher than the visual stimulus presented in isolation. According to the scalar expectancy theory, auditory stimuli may accelerate the pulse rate for visual time perception. Also, the latency of the visual mode switch may be changed by the simultaneously presented auditory stimulus.

Key words: Time perception; Temporal ventriloquism; Auditory-visual integration