2017, Vol. 25, No. 12, 2099–2110
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2017.02099

*

[1] [2] [3] [2] [4] [1]

([1] , 100080) ([2] ( ), 510370) ([3] , 100875) ([4] , 200241)

**摘 要** , ?
, ,
, ,
,

**关键词** ; ; ; ; ; ; ; ; .
**分类号** B842; B845

“ ”
,
,
Cherry (1953)

“ ” ,
(Du, Kong, Wang, & Li, 2011; Schneider, Li, & Daneman, 2007),
? ,
,
,
( )
;

: 2017-04-03
* ( ) ( Z161100002616017)
: , E-mail: liangli@pku.edu.cn

,

## 1 “ ”

,
,
, Cherry (1953)

( ,
unmasking) ,
(fine structure)
(Huang, Xu, Wu, & Li, 2010; Yang et al., 2007)
(Wu, Cao, Wu, & Li, 2013; Wu et al., 2013; Wu, Zheng, Li, Wu et al., 2017; Wu, Zheng, Li, Zhang et al., 2017)
(Wu, Cao et al., 2012; Wu Li et al., 2012; Wu, Zheng, Li, Wu et al. 2017; Yang et al., 2007)
(Huang,

Huang, Chen, Wu, & Li, 2009; Li, Daneman, Qi, & Schneider, 2004; Li, Kong, Wu, & Li, 2013; Wu et al., 2005)

,

(intelligibility) ,

, ,

( , Wu, Zheng, Li, Wu et al., 2017; Wu, Zheng, Li, Zhang et al., 2017; Zheng et al., 2016),

,

( )

, ,

( )

( , Ding & Simon, 2012; Mesgarani & Chang, 2012; Moon et al., 2014; Power, Foxe, Forde, Reily, & Lalor, 2012; Scott & McGettigan, 2013),

(Wu, Zheng, Li, Wu et al., 2017; Wu, Zheng, Li, Zhang et al., 2017; Zhang, Lu, Wu, & Li, 2014; Zheng et al., 2016); ,

(Du, He et al., 2011)

( Wu, Zheng, Li, Zhang et al., 2017; Wu, Zheng, Li, Wu et al., 2017; Zheng et al., 2016)

,

,

,

,

, “ ”

,

## 2

( )

,

, “ ”

? (binding problem)

Treisman Gelade (1980)

(the feature integration theory, FIT)

,

,

(Burwick, 2014; Feldman, 2013; Spence, 2011; Velik, 2012; von der Malsburg, 1999) ,

( )

### 2.1

(tonotopic organization)

“ ” ,

,

,

(Hilbert transform)

(Hilbert, 1912)

(envelope)

(temporal fine structure, TFS) ,

(

Moore, 2008) ,

,

(harmonic structures) (frequency modulation)

(periodic occurrences of noise-like consonants) (Huang et al., 2011) Smith, Delgutte Oxenham (2002) , , , (where) (what) ( Zeng et al., 2004) “ ” , , ,

**2.2** ( ) , , , ( ) (time varying) , “ ” , , , ( ) ( ) , ( ) ,

( ), (precedence effect, Huang et al., 2011; Li & Yue, 2002) (echo threshold), (Li & Yue, 2002) , , (Huang et al., 2011), (Huang et al., 2008, 2009, 2011; Li et al., 2004, 2013; Wu et al., 2005), , , ( ), , , , ( 4.1) , (attribute capture) (Li, Qi, He, Alain, & Schneider, 2005) , , “ ”, ,

**2.3** , , , “ ”, “ ” , , , ( ) “ ” , , “ ” , ,

,

## 3

### 3.1

( )

,

1)

2)

3)

( )

(Li et al., 2004)

,

,

, ,

3 ms (

), ,

3 ms (

)

,

(

),

,

, (perceived spatial separation)

,

(

)

,

(Li et al., 2004; Wu et al., 2005; Rakerd, Aaronson, & Hartmann, 2006; Freyman, Balakrishnan, & Helfer, 2008; Huang et al., 2009; Huang, Wu, & Li, 2009)

,

,

,

/

,

,

,

(Event-Related Potentials, ERP) N1/P2

(Zhang et al., 2014, 2016)

,

N1/P2 (Zhang et al., 2014)

(Du, He et al., 2011)

,

(functional Magnetic Resonance Imaging, fMRI)

($r$ = 0.568, $p$ = 0.009)

,

(Fornito, Yoon, Zalesky, Bullmore, & Carter, 2011; Schulz, Bédard, Czarnecki, & Fan, 2011; Shenhav, Botvinick, & Cohen, 2013) ,

,

( ),

,

(Vouloumanos, Kiehl, Werker, & Liddle, 2001), ,

(Carreiras, Mechelli, Estévez, & Price, 2007)

(Liu et al., 2006)

, “ ”

( Wu, Chen, Wu, & Li, 2014),

“

”

(default mode network, DMN)

(Andrews-Hanna, 2012; Raichle et al., 2001; Whitfield-Gabrieli et al., 2009; Zhang & Li, 2012)

,

(Zheng et al., 2016)

,

,

1)

,

; 2)

; 3) ,

( ) ; 4)

(the motor theory,

Wu et al., 2014),

( )

, ,

,

, (

)

,

**3.2**

“ ” ,

,

“ ”

,

,

,

,

(nonsense sentences), “ ____ ____ ____” ,

(Freyman et al., 2008; Li et al., 2004; Wu et al., 2005; Yang et al., 2007)

,

,

, ( )

( ) ( )

(Gao et al., 2014; Freyman, BalakrishTJ/366.5(s)4.8(hTJ/[ )]TJ6(& Helfer),

; 4)

,

; 5)

( )

,

,

(saliency) ; 6)

( )

; 7)

,

,

fMRI

(Wu, Zheng, Li, Wu et al., 2017)

,

, /

,

/

($r$ = 0.512, $p$ = 0.048)

,

(Ahveninen et al., 2006; Boatman, 2004; Friederici, Rueschemeyer, Hahne, & Fiebach, 2003; Hickok & Poeppel, 2004; Rauschecker & Scott, 2009, 2015; Scott & Wise, 2003)

- (Snijders et al., 2009),

(Papoutsi, Stamatakis, Griffiths, Marslen-Wilson, & Tyler, 2011; Tyler, Wright, Randall, Marslen-Wilson, & Stamatakis, 2010; Tyler, Cheung, Devereux, & Clarke, 2013).

,

(Tong et al., 2005)

,

,

(Arnott, Grady, Hevenor, Graham, & Alain, 2005),

(Lau, Phillips & Poeppel, 2008)

,

/

/

(Ali, Green, Kherif, Devlin, & Price, 2010; Ketteler, Kastrau, Vohn, & Huber, 2008; Li, Yan, Sinha, & Lee, 2008; Menon, Adleman, White, Glover, & Reiss, 2001),

(Thompson-Schill, Bedny, & Goldberg, 2005),

,

,

(Herholz et al., 1996; Papathanassiou et al., 2000; Paulesu et al., 1997; Rodd, Johnsrude, & Davis, 2012; Schuhmann, Schiller, Geobel, & Sack, 2009; Snijders et al., 2009) ,

“ ”

(Wu et al., 2014), /

**3.3** ( )

, ( , visual speech)

,

,

“ ”

(Wu, Cao et al., 2013; Wu, Li et al., 2013; Wu, Zheng, Li, Zhang et al., 2017)

,

,

(Summerfield, 1979)

2010) (Rodd et al.,
, , , 2012) (Romanski, 2012)
“ ” (Giraud & Truy, 2002)
- (temporal synchronization) ,
,
,
, ,

fMRI
, ,
(Wu, Zheng, Li, Zhang et al.,
2017) ,
( )

($r$ = 0.611, $p$ = 0.012) ,

,
(1)
, (2)
, (3)
, (4)

,
(1)
(Campbell et al., 2001; Ludman et al., 2000; Xu, Gannon, Emmorey, Smith, & Braun, 2009); (2)
(Ranganath, 2006; Ranganath, Cohen, Dam, & D'Esposito, 2004; Woloszyn & Sheinberg, 2009); (3)
(Giraud & Truy, 2002; Mummery et al., 1999; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; Wise et al., 1991); (4)
(Chelazzi, Duncan, Miller, & Desimone, 1998; Chelazzi, Miller, Duncanf, & Desimone, 1993; Zhang et al., 2011)
, ,
,
(Alain et al., 2005)
discrimination of speech sounds (Ikeda et al.,

,

**4**

,

,

,

, “ ”

,

, ,

,

: 1)

,

; 2) ,

,

;

3)

,

,

“ ” “ ”

,

, ,

“ ” “ ” ,

“ -

” (1) (2)

(3) (4)

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., .... Belliveau, J. W. (2006).

task control in humans. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 11073–11078.

Du, Y., Kong, L. Z., Wang, Q., Wu, X. H., & Li, L. (2011). Auditory frequency-following response: A neurophysiological measure for studying the "cocktail-party problem". *Neuroscience & Biobehavioral Reviews, 35*, 2046–2057.

Du, Y., He, Y., Ross, B., Bardouille, T., Wu, X. H., Li, L., & Alain, C. (2011). Human auditory cortex activity shows additive effects of spectral and spatial cues during speech segregation. *Cerebral Cortex, 21*, 698–707.

Feldman, J. (2013). The neural binding problem (s). *Cognitive Neurodynamics, 7*, 1–11.

Fornito, A., Yoon, J., Zalesky, A., Bullmore, E. T., & Carter, C. S. (2011). General and specific functional connectivity disturbances in first-episode schizophrenia during cognitive control performance. *Biological Psychiatry, 70*, 64–72.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America, 115*, 2246–2256.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2008). Spatial release from masking with noise-vocoded speech. *The Journal of the Acoustical Society of America, 124*, 1627–1637.

Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex, 13*, 170–177.

Gao, Y. Y., Cao, S. Y., Qu, T. S., Wu, X. H., Li, H. F., Zhang, J. S., & Li, L. (2014). Voice-associated static face image releases speech from informational masking. *PsyCh Journal, 3*, 113–120.

Giraud, A. L., & Truy, E. (2002). The contribution of visual areas to speech comprehension: A PET study in cochlear implants patients and normal-hearing subjects. *Neuropsychologia, 40*, 1562–1569.

Helfer, K. S., & Freyman, R. L. (2009). Lexical and indexical cues in masking by competing speech. *The Journal of the Acoustical Society of America, 125*, 447–456.

Herholz, K., Thiel, A., Wienhard, K., Pietrzyk, U., Von Stockhausen, H. M., Karbe, H., .... Heiss, W. D. (1996). Individual functional anatomy of verb generation. *NeuroImage, 3*, 185–194.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67–99.

Hilbert, D. (1912). *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. Leipzig, Berlin: B. G. Teubner.

Hill, K. T., & Miller, L. M. (2010). Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex, 20*, 538–590.

Huang, Y., Huang, Q., Chen, X., Qu, T. S., Wu, X. H., & Li, L. (2008). Perceptual integration between target speech and target-speech reflection reduces masking for target-speech recognition in younger adults and older adults. *Hearing Research, 244*, 51–65.

Huang, Y., Huang, Q., Chen, X., Wu, X. H., & Li, L. (2009). Transient auditory storage of acoustic details is associated with release of speech from informational masking in reverberant conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1618–1628.

Huang, Y., Li, J. Y., Zou, X. F., Qu, T. S., Wu, X. H., Mao, L. H., .... Li, L. (2011). Perceptual fusion tendency of speech sounds. *Journal of Cognitive Neuroscience, 23*, 1003–1014.

Huang, Y., Wu, X. H., & Li, L. (2009). Detection of the break in interaural correlation is affected by interaural delay, aging, and center frequency. *The Journal of the Acoustical Society of America, 126*, 300–309.

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

Ikeda, Y., Yahata, N., Takahashi, H., Koeda, M., Asai, K., Okubo, Y., & Suzuki, H. (2010). Cerebral activation associated with speech sound discrimination during the diotic listening task: An fMRI study. *Neuroscience Research, 67*, 65–71.

Jeurissen, D., Sack, A. T., Roebroeck, A., Russ, B. E., & Pascual-Leone, A. (2014). TMS affects moral judgment, showing the role of DLPFC and TPJ in cognitive and emotional processing. *Frontiers in Neuroscience, 8*, 18.

Ketteler, D., Kastrau, F., Vohn, R., & Huber, W. (2008). The subcortical role of language processing. High level linguistic features such as ambiguity-resolution and the human brain; an fMRI study. *NeuroImage, 39*, 2002–2009.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De) constructing the N400. *Nature Reviews Neuroscience, 9*, 920–933.

Lesh, T. A., Niendam, T. A., Minzenberg, M. J., & Carter, C. S. (2011). Cognitive control deficits in schizophrenia: Mechanisms and meaning. *Neuropsychopharmacology, 36*, 316–338.

Li, C. S. R., Yan, P. S., Sinha, R., & Lee, T. W. (2008). Subcortical processes of motor response inhibition during a stop signal task. *NeuroImage, 41*, 1352–1363.

Li, H. H., Kong, L. Z., Wu, X. H., & Li, L. (2013). Primitive auditory memory is correlated with spatial unmasking that is based on direct-reflection integration. *PLoS One, 8*, e63106.

Li, L., Daneman, M., Qi, J. G., & Schneider, B. A. (2004). Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *Journal of Experimental Psychology:*

*Human Perception and Performance, 30*, 1077–1091.

Li, L., Qi, J. G., He, Y., Alain, C., & Schneider, B. A. (2005). Attribute capture in the precedence effect for long-duration noise sounds. *Hearing Research, 202*, 235–247.

Li, L., & Yue, Q. (2002). Auditory gating processes and binaural inhibition in the inferior colliculus. *Hearing Research, 168*, 98–109.

Liu, L., Peng, D. L., Ding, G. S., Jin, Z., Zhang, L., Li, K., & Chen, C. S. (2006). Dissociation in the neural basis underlying Chinese tone and vowel production. *NeuroImage, 29*, 515–523.

Ludman, C. N., Lecturer, S., Summerfield, A. Q., Hall, D., Elliott, M., Foster, .... Morris, P. G. (2000). Lip-reading ability and patterns of cortical activation studied using fMRI. *British Journal of Audiology, 34*, 225–230.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain*

*Cortex, 45*, 1111–1116.

Schulz, K. P., Bédard, A. C. V., Czarnecki, R., & Fan, J. (2011). Preparatory activity and connectivity in dorsal anterior cingulate cortex for cognitive control. *NeuroImage, 57*, 242–250.

Scott, S. K., & McGettigan, C. (2013). The neural processing of masked speech. *Hearing Research, 303*, 58–66.

Scott, S. K., & Wise, R. J. (2003). PET and fMRI studies of the neural basis of speech perception. *Speech Communication, 41*, 23–34.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*, 217–240.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature, 416*, 87–90.

Snijders, T. M., Vosse, T., Kempen, G., van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex, 19*, 1493–1503.

Sokol-Hessner, P., Hutcherson, C., Hare, T., & Rangel, A. (2012). Decision value computation in DLPFC and VMPFC adjusts to the available decision time. *European Journal of Neuroscience, 35*, 1065–1074.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*, 971–995.

Summerfield, Q. (1979). Use of visual information for phonetic perception. *Phonetica, 36*, 314–331.

Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology, 15*, 219–224.

Tong, Y. X., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y. S., .... Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage, 28*, 417–428.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Tyler, L. K., Cheung, T. P. L., Devereux, B. J., & Clarke, A. (2013). Syntactic computations in the language network: Characterizing dynamic network properties using representational similarity analysis. *Frontiers in Psychology, 4*, 271.

Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Reorganization of syntactic processing following left-hemisphere brain damage: Does right-hemisphere activity preserve function? *Brain, 133*, 3396–3408.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*, 254–256.

Velik, R. (2012). From simple receptors to complex multimodal percepts: A first global picture on the mechanisms involved in perceptual binding. *Frontiers in Psychology, 3*, 259.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron, 24*, 95–104.

Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience, 13*, 994–1005.

Whitfield-Gabrieli, S., Thermenos, H. W., Milanovic, S., Tsuang, M. T., Faraone, S. V., McCarley, R. W., … Seidman, L. J. (2009). Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 1279–1284.

Wise, R., Chollet, F., Hadar, U. R. I., Friston, K., Hoffner, E., & Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain, 114*, 1803–1817.

Woloszyn, L., & Sheinberg, D. L. (2009). Neural dynamics in inferior temporal cortex during a visual working memory task. *Journal of Neuroscience, 2*9, 5494–5507.

Wu, C., Cao, S. Y., Wu, X. H., & Li, L. (2013). Temporally pre-presented lipreading cues release speech from informational masking. *The Journal of the Acoustical Society of America, 133*, EL281–EL285.

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

Wu, C., Li, H. H., Tian, Q., Wu, X. H., Wang, C. Y., & Li, L. (2013). Disappearance of the unmasking effect of temporally pre-presented lipreading cues on speech recognition in people with chronic schizophrenia. *Schizophrenia Research, 150*, 594–595.

Wu, C., Zheng, Y., Li, J., Wu, H., She, S., Liu, S., Ning, Y., & Li, L. (2017). Brain substrates underlying auditory speech priming in healthy listeners and listeners with schizophrenia. *Psychological Medicine, 47*, 837–852.

Wu, C., Zheng, Y. J., Li, J. H., Zhang, B., Li, R. K., Wu, H. B., … Li, L. (2017). Activation and Functional Connectivity of the Left Inferior Temporal Gyrus during Visual Speech Priming in Healthy Listeners and Listeners with Schizophrenia. *Frontiers in Neuroscience, 11*, 107.

Wu, M. H., Li, H. H., Gao, Y. Y., Lei, M., Teng, X. B., Wu, X. H., .... Li, L. (2012). Adding irrelevant information to the content prime reduces the prime-induced unmasking effect on speech recognition. *Hearing Research, 283*, 136–143.

Wu, X. H., Wang, C., Chen, J., Qu, H. W., Li, W. R., Wu, Y. H., .... Li, L. (2005). The effect of perceived spatial separation on informational masking of Chinese speech.

*Hearing Research, 199*, 1–10.

Wu, Z. M., Chen, M. L., Wu, X. H., & Li, L. (2014). Interaction between auditory and motor systems in speech perception. *Neuroscience Bulletin, 30*, 490–496.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 20664–20669.

Yang, Z. G., Chen, J., Huang, Q., Wu, X. H., Wu, Y. H., Schneider, B. A., & Li, L. (2007). The effect of voice cuing on releasing Chinese speech from informational masking. *Speech Communication, 49*, 892–904.

Zeng, F. G., Nie, K. B., Liu, S., Stickney, G., Del Rio, E., Kong, Y. Y., & Chen, H. B. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America, 116*, 1351–1354.

Zhang, C. X., Arnott, S. R., Rabaglia, C., Avivi-Reich, M., Qi, J., Wu, X. H., … Schneider, B. A. (2016). Attentional modulation of informational masking on early cortical representations of speech signals. *Hearing Research, 331*,