PsyCh Journal
INSTITUTE OF PSYCHOLOGY
Chinese Academy of Sciences

PsyCh Journal **9** (2020): 429–443
DOI: 10.1002/pchj.390

The three second time window in poems and language processing in general: Complementarity of discrete timing and temporal continuity

Xinchi Yu[1,2] and Yan Bao[1,3,4,5]

[1]School of Psychological and Cognitive Sciences, Peking University, Beijing, China, [2]Department of Chinese Language and Literature, Peking University, Beijing, China, [3]Institute of Medical Psychology, Ludwig Maximillian University, Munich, Germany, [4]Human Science Center, Ludwig Maximillian University, Munich, Germany, [5]Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

Abstract: The existence of discrete time windows has triggered the search for permanence and continuity for artists (including poets) in multiple cultures throughout history. In this article, we argue that there exists a 3-s window in the temporal structure of poems as well as in the aesthetic appreciation of poetry by reviewing previous literature on the temporal aspects of poems. This 3-s window can also be considered to be a general temporal machinery underlying human behavior, including language production and perception in general. The reafference principle has provided us a unique frame for understanding cognitive processes. However, "time" was absent in the original two-stage reafference principle. Therefore, we propose a three-stage cycling model of language perception, taking into account time and time windows. We also inspect the possible neural implementations of the three stages: the generation, maintenance, and comparison of predictions (as well as the integration of predictions into the representational context). These three stages are embedded in a temporal window of ~3 s and are repeated in a cycling mode, resulting in the representational context being continuously updated. Thus, it is possible that "semantics" could be carried forward across different time windows, being a "glue" linking the discrete time windows and thus achieving the subjective feeling of temporal continuity. Candidates of such "semantic glue" could include semantic and syntactic structures as well as identity and emotion.

Keywords: language perception; poem appreciation; reafference principle; temporal perception; time window

Correspondence Professor Yan Bao, School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing 100871, China. Email: baoyan@pku.edu.cn

Received 28 April 2020. Accepted 21 June 2020.

Observations of discrete time windows has triggered the search for permanence and continuity for artists and cognitive scientists alike. Reconciling discrete temporal windows and perceptual continuity has always been a challenge in the field of cognitive science. In this article, we first review current evidence for a specific time window, the 3-s window, in the recitation and appreciation of poems as well as in the production and perception of language; then, we propose a three-stage model as a possible basis for perceptual continuity despite the discreteness of time windows, which not only accounts for perceptual continuity in language processing but also potentially in cognitive processing in general.

The pursuit of permanence

East and West, poets in all cultures and at all ages refer to time in their poems. Many poets have emphasized the discreteness of human life; every human owns only one time window of hopefully some tens of years. These time

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

windows come to an end, and all humans are subject to a finite life span, as expressed for instance by the German poet Friedrich Schiller (1759–1805) in “Nänie:” “Auch das Schöne muß sterben! Das Menschen und Götter bezwinget” (“Even the beautiful must perish! It conquers human and gods alike”), or by the Spanish poet Federico García Lorca (1898–1936) in “Alma Ausente:” “No te conoce el niño ni la tarde, porque te has muerto para siempre” (“The child and the afternoon do not know you, because you have died for ever”). The Chinese poet Zhang Ruoxu (660–720) was a pessimist: “人生代代无□已/江月年年只相似” (“Generations and generations, there is no end/Years and years, the moon on the river is just the same”); the only glue of time seems to be the merciless moon.

However, many artists have also tried to overcome such discreteness to create continuity or permanence and even infinity. This pursuit for permanence is a common sentiment across cultures, and it is reflected in different forms of arts. Paintings, photographs, poems, and essays capture a scene and make it permanent as long as the artwork exists. Epic paintings make a moment of an event permanent; scenery paintings make beautiful scenes permanent, and travel notes and diaries keep the episodes alive whenever and wherever someone reads them. The existence of artwork itself marks the permanence of the work and the artist who created it. The artists, although living within the time window of their existence only, create permanence through their works, passing down through generations. This has been long expressed by poets: They know that they will not perish, and will glue the time windows of the human life spans together to achieve continuity and permanence through their works and thoughts. The Roman poet Horace (Quintus Horatius Flaccus, 65 B.C.–8 B.C.) claimed that “Annorum series et fuga temporum./Non omnis moriar multaque pars mei/Vitabit Libitinam: isque ego postera. . .” (I would not entirely die and a large part of me will avoid Libitina [the goddess of funerals and burial]; fresh, I continually would grow with future praise. . .). Similarly, the English poet William Shakespeare (1564–1616) expressed in Sonnet 18 that “So long as men can breathe, or eyes can see,/So long lives this, and this gives life to thee;” “this” being the poem. Poems, stories, and theories (in the world of science) will continue to exist after the time window of one’s own existence ends; later, other people can still pick them up to read, to praise, and to elaborate. The discrete time windows of life cannot only be glued by a merciless moon (as for Zhang Ruoxu in 春江花月夜 “A Moonlit Night On The Spring River”) but also by the cycle of generating and receiving information and knowledge (Pöppel & Bao, 2011). These works of art collectively indicate that the seemingly impermanence and discreteness of human life can become permanent and continuous through information processed and knowledge created that connect generations across ages in history.

A temporal constant of ~3 s in poetry

In addition to the pursuit of permanence, one can identify another temporal characteristic in poetry when poems are recited or read aloud. Poems are composed of verses or *lines* (for a definition of lines, see Turner & Pöppel, 1988). Each line in a poem takes up a certain duration of time. Is there a universal time constant for poetic lines when people read them aloud or recite them by heart? Turner and Pöppel (1988) collected over 20 types of poetry, varying from East to West and from modern societies to indigenous cultures, and found a constant of ~3 s for the duration of poetic lines.

One may argue that this observation is a consequence of the number of syllables, as there is the possibility that every culture may have a similar syllable number per poetic line. Given the temporal constancy of syllable articulation, the observation of ~3 s might be the result of mere item-based (i.e., syllable-number-based) characteristics. We reexamined the data presented by Turner and Pöppel (1988) where the poems have a constant syllabic length; the result is plotted in Figure 1. The mean of the temporal length of lines is 3.09, with an *SD* of 0.64 s, as plotted in this figure (Lines indicate $M \pm SD$.) We did not get a significant Pearson correlation, $p = .348$, $r = 0.261$, nor a Spearman correlation, $p = .314$, $\rho = 0.279$, of these two variables based on the poetry types with constant syllables (which are Japanese epic meter and tanka; Chinese four, five, and seven-syllable line; English seven-syllable trochaic line and octosyllabic ballad meter, ancient Greek dactylic hexameter half-line, trochaic tetrameter half-line, iambic trimester, anapestic tetrameter half-line, Latin elegiac couplet, dactylic hexameter half-line, hendecasyllabic, and French 12-syllabic alexandrine; see Table 1). Note that a significant correlation is still not present after removing the data point of “Latin elegiac couplet.” Bearing the risk of interpreting null results in statistics, this result suggests that the

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

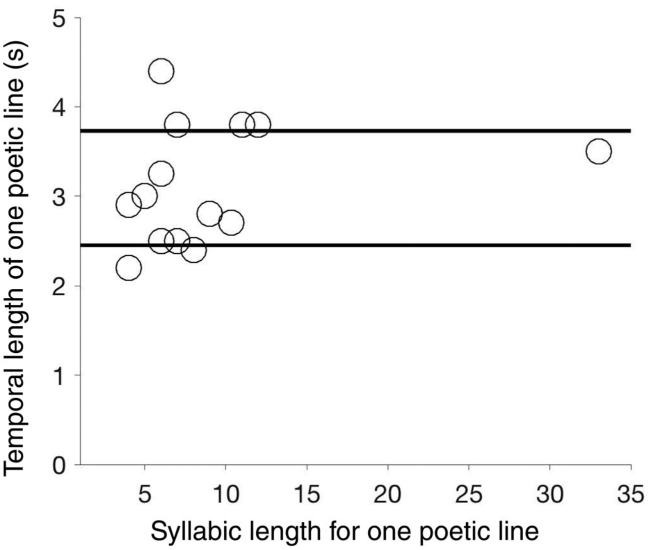

Figure 1. Visualizing syllabic length and temporal length of one line in different type of poems from Turner and Pöppel (1988).

temporal constant of poetic lines is not a result of a constancy in syllabic length but presumably a temporal constant.

Our conclusion fits with our experience that we automatically speak slower with less syllables in one poetic line, and relatively faster if there are more syllables; if there are too many syllables in one "sentence," we will enter a pause in the middle and separate them into multiple lines in recitation (Pöppel, 1988). One typical example of slowing down comes from the poem of "Il pleure dans mon cœur" by the French poet Paul Verlaine (1844–1896): "C'est bien la pire peine/De ne savoir pourquoi/Sans amour et sans haine,/Mon cœur a tant de peine." (And the worst pain of all/Must be not to know why/Without love and without hate/My heart feels such pain (Richard Stokes, Trans.). In this example, one automatically speaks slower when there are fewer syllables in one line. Less words are employed in one line to express a depression in this case; the pace of citing is slowed down to fit into the 3-s time window.

Table 1
Syllabic Length and Temporal Length of Poem Types with Constant Syllable Number from Turner and Pöppel *(*1988*)*

Poem type	Average syllabic length per line	Average temporal length per line(s)
Japanese epic meter	6	3.25
Japanese Tanka	10.33	2.70
Chinese four-syllable line	4	2.20
Chinese five-syllable line	5	3.00
Chinese seven-syllable line	7	3.80
English seven-syllable trochaic line	7	2.50
English ballad meter (octosyllabic)	8	2.40
Ancient Greek dactylic hexameter (half-line)	9	2.80
Ancient Greek trochaic tetrameter (half-line)	4	2.90
Ancient Greek iambic trimeter	6	4.40
Ancient Greek anapestic tetrameter (half-line)	6	2.50
Latin elegiac couplet	33	3.50
Latin dactylic hexameter (half-line)	9	2.80
Latin hendecasyllabic	11	3.80
French alexandrine (12-syllable)	12	3.80

There was, however, severe criticism of Turner and Pöppel (1988) by Fabb (2013), who claimed that the 3 s constant in poetic lines does not exist. Fabb made this conclusion mainly based on a recording corpus of English poems that he collected and argued that there are plenty of poetic lines exceeding 3 s. There may have been, however, a misunderstanding by Fabb concerning the 3-s time window; it is not a window with a ceiling of exactly 3 s as one would expect from a physical constant but is an operational platform based on statistical analyses reflecting an endogenous neural mechanism with some variance being typical for biological processes. This means that many poetic lines have a duration of ~3 s, with some variance to give rise to a median/mean value of ~3 s. This mean value can be shown also in Fabb's own corpus. Fabb (2013) recorded 55 English poems, and the data of the median/mean line duration for *each poem* were listed in the article. In Figure 2, we plot the distribution of the mean and the median of line duration in Fabb's article. The mean of median duration (for each poem) is 3.28 s, SD = 0.853 s, and the mean of mean duration is 3.38 s, SD = 0.893 s. As two slightly skewed distributions (skewness of median duration = 0.240, skewness of mean duration is 0.146, note that the two distributions all pass a Shapiro–Wilk normality test, $p > .05$ whereas the distribution of median duration exhibits a $p = .049$ in the Kolmogorov–Smirnov test with Lilliefors correction), the median of median duration is 3.10 s, and the median of mean duration is 3.22 s. This result, in contrast to what Fabb argued, actually shows the robustness of the 3-s time window, as it applies to a new set of poems collected by another independent researcher. Exceeding the exact value of ~3 s does not invalidate the existence of this time window in poems; it is not unusual

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

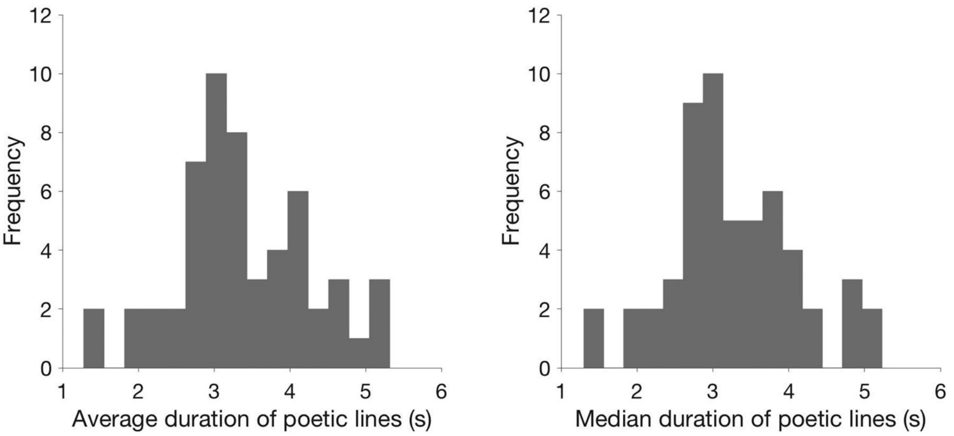

Figure 2. The distribution of the temporal length of poetic lines, based on data from Fabb (2013).

that human cognitive processes are subject to variation (e.g., Bao, Yang, Lin, & Pöppel, 2016).

One might still argue that this notion of a 3-s time window is trivial, as it matches approximately the normal human respiratory cycle of 12 to 15 breaths per minute (Barrett, Barman, Boitano, & Brooks, 2012). One reason for this explanation to be unlikely is that poems as works of art do not have any intrinsic characteristics to match the human respiratory cycles. Furthermore, humans are not forced to obey the respiratory cycle in speech (Parkes, 2006). In this context, it is useful to take a more general biological perspective comparing the 3-s window in humans with a potential equivalent in other species. It has been observed that the 3-s window is also a frame for (repetitive) behavioral segments in humans and other higher mammals (Gerstner & Cianfarani, 1998; Gerstner & Fazio, 1995; Gerstner & Goldberg, 1994; Kien, Schleidt, & Schöttner, 1991; Schleidt, Eibl-Eibesfeldt, & Pöppel, 1987). This time window being the same in speech and other motor activities in different species indicates that they share a common logistic temporal mechanism (Pöppel, 1997a). However, the breathing rate of higher mammals does not necessarily match with 3 s. For example, the median duration of motion segments for a giraffe is 2.61 s and the mean is 3.61 s (Gerstner & Fazio, 1995) whereas the mean respiratory rate is 6.57 s (Langman, Bamford, & Maloiy, 1982). Therefore, the pattern of 3 s in speech and other motor behaviors is not a passive reflection of breathing cycles, although we do not dismiss the possibility that respiratory cycles can modulate the duration of such time windows.

Another reason that the 3-s window does not reflect a trivial phenomenon is that poem appreciation is not only productional but also perceptual, thus providing an important synchrony of production and perception on the aesthetic level. By simply altering the temporal length of poetic line (without changing the syllable number), Zhao, Zhang, and Bao (2018) discovered a cross-culturally robust preference of poetic lines in the 3-s range. Even if the subjects cannot understand the meaning of the poetic lines (in this case, German subjects listening to Chinese poems), they still prefer lines of ~3 s. This observation indicates a production–perception synchrony of ~3 s, which means that we not only tend to recite poems (and speech in a more general sense; discussed later) with a 3-s pattern but also appreciate poems aesthetically within the same temporal frame. This temporal preference for a 3-s pattern and not a 1-s or a 10-s pattern, which linguistically would be possible, indicates presumably a profound evolutionary basis. The temporal modulation effect of the 3-s window on aesthetic appreciation may also motivate to look for other concepts and phenomena of the cognitive and neural basis of aesthetic perception in general and in detail, as has been partly already done for decision processes, the visual arts, and music (Avram et al., 2013; Bao et al., 2016; Bao et al., 2017; Park et al., 2014; Park et al., 2015; Pöppel, 1989a).

A necessary distinction between content and logistic functions

Speech is of a transient nature, which obviously means that what has been said no longer physically exists after it has been said; this implies that one has to utilize a segmentation process to store and process information (Christiansen & Chater, 2016) and that this segmentation process most likely has its limits. Some have addressed the nature of such limits as item-based, like the magical number of 7 ± 2 by Miller (1956) or 4 ± 1 by Cowan (2001). Others have addressed the nature of such segmentation as time-based, as suggested by the 3-s time window by Pöppel (1997a), which is the main focus of this article. However, no matter if such segmentation is item-based or time-based, one has to stress that this segmentation process is logistic. This means that segmentation itself is relatively independent of the content that is segmented. No matter what one is segmenting, the segmentation limit is always 7 ± 2 or 4 ± 1 if item-based or ~3 s if time-based, indicating that this limit has a logistic characteristic or is "pre-semantic" (Pöppel, 1997a, 2009). We avoid using the ambiguous term *chunk*, as chunking can happen on multiple levels (Christiansen & Chater, 2016); one has to chunk at least one whole sentence as one to understand it.

The logistic-content distinction appears to be necessary for a deeper understanding of cognitive processes. It is not controversial that percepts, memories, emotions, or volitions represent the content of our cognitive systems. However, temporal organization, attentional control, and general activation or arousal are not on the same categorical level as the content functions. The temporal organization of different content functions cannot be classified at the same logical level as can content functions themselves. Furthermore, attentional systems operate *on* content functions rather than *with* them in parallel (Bao et al., 2011; Bao & Pöppel, 2007; B. Zhou, Bao, Sander, Trahms, & Pöppel, 2010). Thus, attentional processes can also not be classified at the same logical level as content functions. In addition, arousal modulated by the activation system represents another neural process that is logically different from content functions; it is functioning like a "power supply" of the central nervous system. Therefore, it is necessary to separate the logistic functions (temporal control, attentional mechanisms, and activation of functions) from the content functions. The logistic functions are the "how" functions providing the necessary neuronal infrastructure for "the state of being conscious" (Pöppel, 1997b) whereas the content functions are the "what" functions that are related to conscious representation (Pöppel, 1989b; B. Zhou et al., 2016).

Such logistic-content distinction can be seen in physiological and imaging studies of auditory speech processing. For example, the acoustic analysis of the auditory temporal structure is observed in the superior temporal sulcus whereas speech-related activations appear in more brain areas (Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Overath, McDermott, Zarate, & Poeppel, 2015). Following these findings, we can conjecture that the logistic "containers" root partly in the functionally or temporally earlier areas in the processing routes. For example, bistable perception of the Necker cube can be decoded even in early visual areas in the posterior visual cortex (M. Wang, Arteaga, & He, 2013). The concept of central timing mechanisms has been discussed for quite some time, based for instance on the similar temporal order threshold of tens of milliseconds or the multimodal response distributions of reaction time and eye movement latencies, and the low-frequency time window of several seconds across modalities (Hirsh & Sherrick, 1961; Pöppel, 1970, 1994, 1997a; Pöppel & Logothetis, 1986; Wittmann, 1999). Taken together, the distinction between content and logistic functions appears to be necessary, reflecting fundamental biological mechanisms, and temporal processing serves as one of these logistic functions.

A logistic time window of ~3 s in language

The 3-s synchrony in the recitation and appreciation of poetry is similarly observed in language in general: Linguistic utterances have the tendency to be expressed with temporal segments of ~3 s (Kien & Kemp, 1994; Kowal, O'Connell, & Sabin, 1975; Schleidt & Kien, 1997; Vollrath, Kazenwadel, & Krüger, 1992; note that the results of Kowal et al., 1975) were re-analyzed for more clarity; see Table 2). These findings indicate that there is a periodic nature of ~3 s in human verbal behavior, meaning that a conspicuous pause often appears approximately every 3 s. We argue that besides being a logistic segmentation window for language production, the logistic 3-s time window is also underlying language perception. Note that in this context, we do not distinguish systematically between language perception and language processing (as in Hickok & Poeppel, 2007; note that in the citations of this article, "Poeppel" refers to Dr. David Poeppel, and "Pöppel" refers to Dr. Ernst Pöppel).

Both behavioral and electrophysiological evidence for other cognitive functions support the concept that the interval of 3 s is something special in temporal processing (Bao et al., 2013; Elbert, Ulrich, Rockstroh, & Lutzenberger, 1991; Fraisse, 1984; Kagerer, Wittmann, Szelag, & Steinbüchel, 2002; Pöppel, 1972, 1997a). A 3-s time

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

Table 2
Average Duration per Speech Segment *(*re-analysis based on Kowal et al., 1975*)*

Age	Syllables per second (Kowal et al., 1975)	Syllables per speech segment (Kowal et al., 1975)	Average duration per speech segment (s)
Kindergarten	2.15 (0.75)	5.5 (1.9)	2.56
Grade 2	2.86 (0.53)	7.4 (2.2)	2.59
Grade 4	3.24 (0.51)	7.6 (2.8)	2.35
Grade 6	3.26 (0.66)	7.3 (2.5)	2.24
Grade 8	3.83 (0.50)	9.4 (2.4)	2.45
Sophomores	4.00 (0.51)	9.9 (3.2)	2.48
Seniors	3.84 (0.52)	10.0 (3.2)	2.60

Note. Numbers in parentheses represent *SD*s. The average duration per speech segment is calculated from syllables per speech segment divided by syllables per second.

window has for instance been demonstrated in experiments on auditory bistable perception, indicating perceptual segmentation. When listening to ambiguous (i.e., bistable) auditory stimuli such as “KU-BA-KU,” people switch their perception spontaneously approximately every 3 s between “KUBA” and “BAKU” (Radilová, Pöppel, & Ilmberger, 1990; von Steinbüchel, 1998). This phenomenon of perceptual alteration applies also for bistable visual figures (Gómez, Argandoña, Solier, Angulo, & Vázquez, 1995; Polgári, Causin, Weiner, Bertschy, & Giersch, 2020; von Steinbüchel, 1998). A 3-s time window in auditory temporal segmentation is also indicated electrophysiologically by the event-related potential (ERP) component of mismatch negativity (MMN; L. Wang et al., 2016; L. Wang, Lin, Zhou, Pöppel, & Bao, 2015) and magnetoencephalography experiments (Sams, Hari, Rif, & Knuutila, 1993). A larger MMN for deviant stimuli is elicited if the interval between successive auditory stimuli is around 3 s; the MMN amplitude decreases after 3 s, but restores again after around 6 s, indicating an oscillatory neural mechanism with a period of ~3 s. The fact that the human cognitive system is more open for new stimuli approximately every 3 s is sometimes misunderstood as by White (2017); he claimed that the MMN being observed also after an interval of 9 to 12 s would contradict the hypothesis of temporal integration on a time scale of ~3 s; this is certainly a misunderstanding. Note, however, that the concept of a 3-s time window or a high-frequency time window of ~30 ms to 40 ms does not exclude time windows between these periods; for example, the window of 150 ms to 300 ms (e.g., Ding et al., 2017; Hickok & Poeppel, 2007; Poeppel & Assaneo, 2020), in which suprasegmental information is carried. This inclusion of other temporal windows has actually already been stated by Szelag and Pöppel (2000).

Abundant behavioral and neurophysiological evidence has proven the existence of a 3-s time window also in language perception. Early behavioral evidence came from Sachs (1974), where subjects were asked to report if a sentence was identical to a previous one. The two sentences were distanced by sentences with different numbers of syllables. There is a significant drop and a plateau after the 3-s window both for auditory stimuli (between 3 s and 7.5 s) and visual stimuli (between 1 s and 4 s). Such results suggest that subjects in these experiments were more sensitive to changes in a sentence if they are in the same temporal segment of ~3 s. Wagers and Phillips (2009) discovered that the plausibility effect of a sentence is weaker for longer dependencies than are shorter ones in a self-paced reading task. More significant differences in reading time after the critical verb onset was observed for short dependency conditions than for long dependency ones. In short dependency conditions, the distance; that is, the interstimulus interval (ISI), between the sentence-initial noun and the critical verb is 1 (~450 ms, as estimated from the figure as the original data were not provided) whereas the distance is 6 (~3 s) in long dependency conditions. Thus, a difference within and beyond the 3-s window is observed in this experiment. One may argue that more interval conditions are needed to confirm that this effect is truly a staircase effect (indicating a window) rather than a simple linear decreasing effect. Nonetheless, Wagers and Phillips (2009) provided some positive implications about the existence of this time window. Also note that such findings are unlikely to be a trivial result of the human memory limits of item (i.e., 7 ± 2 or 4 ± 1), as one needs to chunk the whole sentence holistically in language perception, and there was no significant difference in answering the follow-up questions, indicating that the subjects' memory can equally hold up short and long sentences in this experimental condition. Schremm, Horne, and Rolls (2015) result has indicated that the reaction time to grammatical and semantic incongruency can be modulated by a 3-s window. Subjects reacted slower if the critical words had a distance longer than ~3 s when the number of words between them was the same.

Electrophysiological evidence also has supported the existence of a 3-s time window in language perception. Strong evidence has come from Roll, Lindgren, Alter, and Horne (2012), who examined the ERP component of the closure positive shift (CPS), a centroparietal positivity related to perceptual boundaries in language and music perception. A more salient CPS effect was observed only when the clause boundaries coincided with the temporal boundary of ~2.7 s. This indicates that there is a logistic wrap-up effect enhancing the wrap-up effect of content. Roll, Gosselke, Lindgren, and Horne (2013) studied the component of the left anterior negativity (LAN) elicited by the anomaly of grammatical agreement. They found that the LAN effect holds for a (SOA) of 1.75 s and 2.25 s, but is not significant for an SOA of 2.75 s. For an SOA beyond the 3-s window (3.25 s), an anterior negativity with different scalp topography was observed, indicating different neural mechanisms inside and outside the 3-s window. Phillips, Kazanina, and Abada (2005) discovered that longer filler-gap dependency (ISI = 4000 ms) will result in a delayed P600 effect, as compared to a shorter dependency (ISI = 2000 ms). A similar experiment with ISIs of 700 ms and 2600 ms did not report such a result on the verb position (Fiebach, Schlesewsky, & Friederici, 2002). This can be explained by the 3-s time window, as an ISI of 700 ms, 2000 ms, and 2600 ms falls into the 3-s window whereas an ISI of 4000 ms does not, thus resulting in a qualitative difference for the P600 latency.

Synchronization in verbal communication

Based on these many different observations, one can conclude that the 3-s window acts as a segmentation frame for both language production and perception. This production–perception synchrony can be considered as the basis for synchronization in verbal communications because it allows the creation of shared temporal windows of communicators. They can synchronize their mutual "nowness" or "subjective present" and share a temporal frame of reference in social interactions (Bao et al., 2015; Miyake, 2012; Nagy, 2011; Pöppel, 1988, 1994, 1997b, 2009).

Synchronization on the basis of a 3-s time frame can account for many aspects of human linguistic behaviors (e.g., in conversations). Sometimes one does not finish while speaking a sentence or omits parts of a sentence; despite this, one can still successfully transmit a verbal message. The fact that human speech segments follow a 3-s time frame (Kowal et al., 1975; Vollrath et al., 1992) also explains that one makes deletions to fit utterances into this 3-s time frame. The implicit consensus that speakers and listeners all operate within a 3-s basis also saves energy, as one does not have to shift between different temporal frames from moment to moment. Thus, synchronization is also the operative basis for anticipation; one implicitly expects a speech segment to stop at around 3 s, and another topic in conversation can be initiated after this time.

This 3-s temporal production–perception synchrony may be based on similar neural mechanisms for speech production and the perception systems as a result of an evolutionary co-adaptation of the time windows. The basic idea of a social co-adaptation is that humans synchronize to each other during numerous social interactions, adjusting individual temporal windows to a common length (here, of ~3 s). Such synchronization in social interactions might involve synchronization on both behavioral and physiological levels (e.g., Dikker et al., 2017; Hale, Ward, Buccheri, Oliver, & Hamilton, 2020). The neural and functional relationship between speech production and perception is not a new topic; it has been discussed with respect to the perception–production functions of Wernicke's area (Binder, 2015) and the motor theory of speech perception (Liberman & Mattingly, 1985). Such low-frequency auditory–motor interaction might also involve low-frequency neural oscillatory coupling (e.g., with delta waves); a similar scenario has been observed for the syllabic level (Assaneo & Poeppel, 2018).

The continuity of 3-s windows can be modeled with oscillations. One has, however, to distinguish between two different classes: pendulum oscillations and relaxation oscillations. The continuity of 3-s windows can technically be described as a pendulum oscillation, and physiologically it can perhaps be related to the delta band wave in the electroencephalogram. The frequency of the delta band (0.5–4 Hz) perhaps matches the cycle of such a window, and there is evidence that the delta band is related to the

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

processing of intonational phrase boundaries (e.g., Bourguignon et al., 2013; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; for a review, see Meyer, 2018), which often fall into the window of ~3 s. However, the 3-s time window can also be modeled by relaxation oscillations (Pöppel, 1970, 1994) because the ongoing pendulum oscillation can be interrupted at any phase, and a new oscillatory process with the same or similar period is entrained with very short latency, a characteristic of relaxation oscillations (for a recent discussion on exogenous and endogenous cortical rhythms in language processing, see Meyer, Sun, & Martin, 2019). Neural and behavioral phenomenon below 1 Hz are often difficult to identify (H. Zhou, Melloni, Poeppel, & Ding, 2016), but they certainly account for a great part of human behavior (e.g., Donhauser & Baillet, 2019; Teng et al., 2020). Individual variance of the period of an oscillatory process can possibly mask an oscillatory phenomenon below 1 Hz, as seen in studies by L. Wang et al. (2016; L. Wang et al., 2015). Therefore, single-case studies can be extremely useful in elucidating such low-frequency oscillatory phenomena (Bao et al., 2015; Bao et al., 2017, c; Pöppel & Bao, 2017; Zaytseva & Bao, 2015).

A three-stage cycling model in language processing

The concept of the reafference principle (von Holst & Mittelstaedt, 1950, 1971) roots from the exploration of the stability of vision by von Helmholtz (1896), and it was later generalized using the term *corollary discharge* (Teuber, 1960). The later notion of “predictive coding” could also be seen as inspired by the reafference principle (e.g., Clark, 2013; Rao & Ballard, 1999; Stefanics, Kremláček, & Czigler, 2014). In the original version of the reafference principle, an efference copy is generated in parallel to an action command for whatever one is doing. This copy is then compared with the sensory information after the execution of the action (i.e., the reafference). Thus, there are two successive operations in the original reafference principle: (a) the generation of an efference copy and (b) the process of reafference allowing the comparison between the executed movement and the efference copy. What is missing in the original concept and the derived theories is the factor “time” (i.e., how long the duration of an efference copy has to be maintained). Thus, the reafference principle should actually comprise three components: (a) generation of an efference copy (which serves as the basis for prediction) in parallel to the motor command, (b) maintenance of the efference copy or the prediction in a temporal buffer, and (c) comparison of the prediction with the sensory input after a specific temporal interval. Another cycle with all its components then can be initiated, be it either periodically under endogenous temporal control or in a voluntarily determined active mode.

A new cycle will not result in a simple repetition of what is represented because there is always an updating in each cycle, which will result each time in modified predictions (Pöppel & Bao, 2011). As has been suggested, many cycles could operate simultaneously in a hierarchical system, such as the higher frequency cycles being embedded in the lower frequency cycles (Tanida & Pöppel, 2006). This theoretical concept provides an explanation for planning and confirming what has been planned as a basis for the continuity of action and perception. This can also be considered as a framework in realizing the original attempt of von Helmholtz (1896) in discussing the nature of visual continuity despite saccadic eye movements (for a review, see Wurtz, 2018) and visual stability across different areas of the visual field which are characterized by different mechanisms of spatial attention (Lei, Bao, Wang, & Gutyrchik, 2012; Pöppel & Bao, 2012).

The reafference principle has indeed been suggested as a theoretical frame for perceptual processes in general, and has been suggested to explain “pre-emptive perception” (Bodis-Wollner, 2008), “proactive perception” (Rimmele, Morillon, Poeppel, & Arnal, 2018), the discontinuity of thinking in schizophrenia (Feinberg, 1978), and speech production (Tian & Poeppel, 2010, 2013; Tian, Zarate, & Poeppel, 2016). Although the reafference principle has been invited into the field of language perception (Gambi & Pickering, 2017; Pickering & Clark, 2014; Pickering & Gambi, 2018; Pickering & Garrod, 2013), the reasoning is basically on the computational and algorithmic level (Marr, 1982). Neurobiological features of the principle refer mainly to perception–production connections (Dick & Andric, 2013; Pickering & Garrod, 2013) rather than to the basis of the two (or three, as we propose) operations in the reafference principle. Pickering and Gambi (2018) introduced a cognitive model of reafference in language perception, but did not consider its neural implementation. We suggest possible neural mechanisms (i.e., the hardware implementations) of reafference in

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

language processing, in poetry in particular and cognitive processes in general, while being compatible with current linguistic concepts.

When processing a sentence, predictions can be generated before the "critical" word actually appears. For example, when we hear/see the first part of a sentence (e.g., You never forget how to ride. . . .), the following part can be predicted (e.g., a bicycle, an elephant, etc.). Predictions can have different weights. Here, "a bicycle" is more likely to appear, but if "an elephant" appears, we know the sentence is still congruent, but more integration efforts should be implemented because "an elephant" is less likely to appear here. Such a notion of prediction and integration, however, is abstract, and it is difficult to be mapped directly onto neural activities in the brain.

We propose that our three-stage model (as outlined earlier) is compatible with language perception. These three stages are not just a theoretical construct but are validated by neural evidence and can be mapped onto certain neural underpinnings. In our framework, the prediction-integration process consists of three consecutive stages (see Figure 3): (a) the generation of predictions, (b) the maintenance of predictions, and (c) the comparison between prediction and the current input (as well as the integration of the current input into the representational context). Note that comparison is not the end of the story: After comparison, the input word is integrated into the current semantic context ("narrative") and syntactic context (i.e., hierarchical structure), which, along with the predictions, has also been persisting through time. To be more precise, we would refer to the "context" as a "gestalt." Therefore, the whole process should also include the generation and maintenance of context, and the integration of the current information into context; this process apparently falls into a larger temporal window. Note that the "integration" we use here can be different from the way it is used in some psycholinguistic research. One prominent feature of our model is that this three-stage cycle operates in a 3-s time window. After one

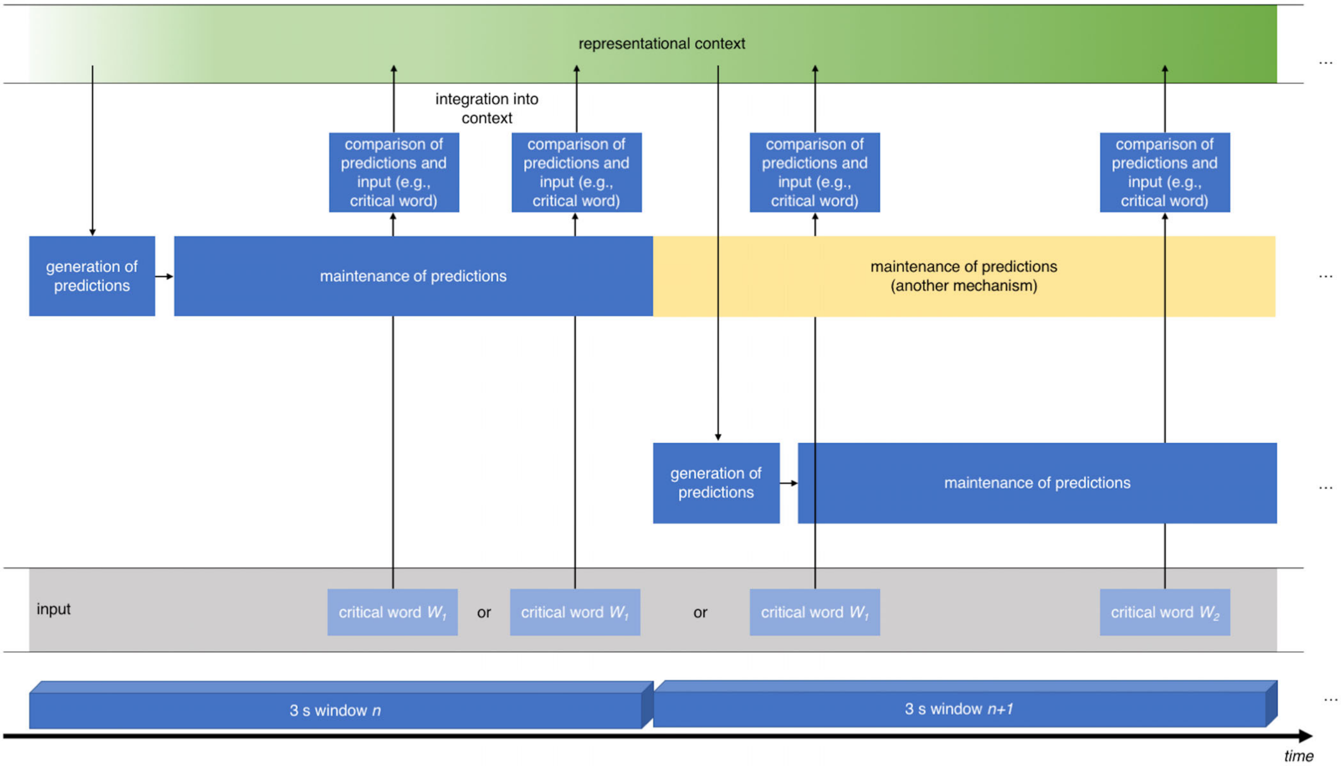

Figure 3. An illustration of the three-stage cycling model in language processing. Note that the change from blue to yellow in the "maintenance of predictions" stage indicates a shift of processing machinery within and outside the 3-s window. Critical word W1 is compared and integrated with predictions maintained by different mechanisms, based on its temporal onset (within/beyond the same 3-s window). If W1 is presented after the previous 3-s window and related to the predictions of the former 3-s window, then it is compared and integrated in a different machinery, as the brain is currently in another 3-s window, generating new predictions, and looking for new information. The cycling three-stage process as well as the continuously updated representational context ("semantics") allow for a temporal continuity across discrete time windows. [Color figure can be viewed at wileyonlinelibrary.com]

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

cycle, the brain actively “looks for” something new, with the launching of a new cycle, which has a range of ~3 s (L. Wang et al., 2016; L. Wang et al., 2015). However, this does not mean that the information maintained in the previous temporal window is lost. Rather, it is taken over by a different cognitive machinery (discussed earlier, in the “A logistic time window of ~3 seconds in language” section), indicated by behavioral and electrophysiological studies (Fiebach et al., 2002; Phillips et al., 2005; Roll et al., 2013; Sachs, 1974; Schremm et al., 2015; Wagers & Phillips, 2009).

Predictions are generated before the “critical word” actually appears, as the brain activities predicting semantically similar words are similar (L. Wang, Kuperberg, & Jensen, 2018). L. Wang et al. (2018) also demonstrated that such predictions are likely to be generated from left inferior/medial temporal lobe. It often takes some time after predictions for a certain word to be generated before it actually appears; therefore, the prediction needs to be held for a certain time. The cognitive demand of maintaining predictions cannot be done without the continuity of specific neural activities. A possible candidate for such neural persistence is the sustained anterior negativity (e.g., Kluender & Kutas, 1993; Lau, 2018; Lau & Liao, 2017; Phillips et al., 2005). During this sustaining process, knowledge from long-term memory could be retrieved to narrow down and re-weight the saliency of predictions, known as the notion of “prediction as memory retrieval” (Chow, Momma, Smith, Lau, & Phillips, 2016). The third stage is comparing the current input with the predictions; such comparison relies on the decodability of neural activities (e.g., de Lange, Heilbron, & Kok, 2018), making it possible to compare “new” percepts with the “old” predictions which are maintained. The new percepts are then integrated into the current gestalt (context). The well-known evoked potential components in some psycholinguistics studies (e.g., N400, P600, etc.) are likely to reflect the third stage due to its temporal characteristics; it appears after the critical word onset (e.g., Kutas & Hillyard, 1980; Zhang, Yu, & Boland, 2010). Every time each three-stage cycle ends, offline knowledge (i.e., long-term memory or experience) is updated and can be utilized in future predictions. In this way, “online” information becomes “offline” and gets stored for future use.

In terms of oscillation, we conjecture that the critical word for prediction (e.g., “ride” given in the example) re-entrains a relaxation oscillation and initiates a new time window; the oscillatory process continues, “waiting” for an appropriate predictive operation (e.g., coupling) until the target word (e.g., “a bicycle”) appears. Thus in this model, the 3-s time window being implemented by successive periods of an oscillation plays the decisive role for understanding.

Continuity across discrete time windows to form discourse gestalts

One might ask an obvious question: If information processing is framed by successive discrete 3-s time windows, why does one not sense the “bumps” or interruptions when entering a new time window? This seemingly killer question can be resolved after our reasonings on the 3-s window and the reafference principle. Predictions do not vanish to nowhere at the end of the 3-s windows. The discrete time window does not stop the predictions from existing; information from the previous time window continues to exist as predictions for the following time windows, and through comparison and integration, the new inputs in the recent time window are connected with information of the previous time window(s). Thus, continuity can be achieved across time windows. Through this process, utterance segments (with a duration of ~3 s) can be “glued” temporally through prediction, comparison, and integration to form a coherent discourse. The reafference principle can function as a framework in understanding how discrete perceptual units through time can be perceived as a holistic gestalt. Contents in every single temporal window are integrated into a gestalt. Such persisting gestalts form the prediction of future input. The predictions can be compared and integrated with future inputs or future gestalts, resulting in a “full” gestalt or a new gestalt leading to future predictions. The three-stage process is used to different extent across individuals, which can be validated by the fact that different people have different abilities in discourse comprehension, and by experiments on holistic and analytical processing strategies (Bao et al., 2013).

The discrete windows are “glued” together because the content (“semantics” in a broader sense, including semantic and syntactic structures) is carried over from previous time windows. Thus, semantics is the glue for the sequence of time windows. Beyond semantic and syntactic structures, such glues are functional in our model, given two presuppositions: the concept of identity “of something as something" and the time constants in the representation of emotions. Identity is the defining characteristic by which a percept, a thought, or a mental image is recognized as a persisting entity over time (Pöppel, 2010; B. Zhou, Pöppel, & Bao, 2014). Identity can be the sameness of an object (or objects); for example, the computer screen (or the paper) one is staring at in the previous time window is still the same in the current time window. Identity also can refer to mental images created for instance by sequential lines in a poem (see the example of the poem by Verlaine). Thus, the maintenance of identity of a percept or a mental image is the necessary (but not sufficient) condition for the creation of a meaningful permanence. The second presupposition in our model to glue time windows together is related to emotional or evaluative neural representations. Emotional states normally last longer and can have a duration of many seconds, minutes, hours, or even days (Verduyn & Lavrijsen, 2015). On that basis, the neural representation of emotional evaluations could act as a functional glue that bridges the sequence of 3-s time windows; in that case, the process of connecting could stay at an implicit level (Pöppel & Bao, 2011). The long-term emotional tonality would provide the frame for continuity. Failure in using semantics as a temporal glue might account for subjective discontinuities, as in schizophrenic patients, for example. Such patients often exhibit discontinuities in their mental activity, and it has been suggested that the disturbance of the semantic connection of successive units could be the reason for this problem (Bleuler, 1969; Martin et al., 2014; Pöppel, 1989b).

Conclusion and summary

As has been stated, behavioral as well as neural evidence support the concept that cognitive functions operate with multiple time windows (Hickok & Poeppel, 2007; Martin, 1972; Szelag & Pöppel, 2000; L. Wang et al., 2016; L. Wang et al, 2015; for an earlier review, see Pöppel, 1997a). We have argued further on the basis of experimental evidence that there is indeed a 3-s time window in poetry and other linguistic behaviors, and that these discrete windows do not disprove perceptual and conceptual continuity. To illustrate the latter point, we introduced the reafference principle as a concept and argued that the three-stage process creates the basis for continuity across temporal windows. This reasoning brings us back to the initial remarks; although one owns only one time window in history, stories, poems, or theories will be passed down through generations. The continuity across discrete windows is achieved through the complementarity as a generative principle of passing and receiving. Thus, our model on continuity of windows can also be seen as being inspired by human anthropological universals (Bao & Pöppel, 2012) from different poets in different windows and different cultures. Not only can the principle of reafference bring about continuity but it also can bring about a holistic gestalt in perception, as has been discussed in discourse comprehension.

We also provided the possible neural and cognitive bases for the successful operation of reafference in language (and possibly other cognitive processes). To further elucidate the process of reafference in language, we also need to settle the underpinnings for reafference on the neurobiological level in future research. It is also worth researching whether different levels of linguistic stimuli (e.g., syllable level, word level, etc.; word category level, semantic level, syntax level, etc.) operate on the same or different neurobiological routes. The reafference principle is also a demonstration of complementarity. Therefore, our reasoning also demonstrates the power of complementarity as a generative principle (Bao, von Stosch, Park, & Pöppel, 2017). Finally, note that poems from different cultures being characterized by anthropological universals provide an ideal data source to reflect on discrete time and temporal permanence.

Acknowledgments

This work was supported by National Natural Science Foundation of China Projects 31771213 and 31371018 to Yan Bao.

Disclosure of conflict of interest

The authors declare that there are no conflicts of interest.

References

Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, *4*(2), eaao3842. https://doi.org/10.1126/sciadv.aao3842

Avram, M., Gutyrchik, E., Bao, Y., Pöppel, E., Reiser, M., & Blautzik, J. (2013). Neurofunctional correlates of esthetic and moral judgments. *Neuroscience Letters*, *534*, 128–132. https://doi.org/10.1016/j.neulet.2012.11.053

Bao, Y., & Pöppel, E. (2007). Two spatially separated attention systems in the visual field: Evidence from inhibition of return. *Cognitive Processing*, *8*(1), 37–44. https://doi.org/10.1007/s10339-006-0151-x

Bao, Y., & Pöppel, E. (2012). Anthropological universals and cultural specifics: Conceptual and methodological challenges in cultural neuroscience. *Neuroscience and Biobehavioral Reviews*, *36*(9), 2143–2146. https://doi.org/10.1016/j.neubiorev.2012.06.008

Bao, Y., Pöppel, E., & Zaytseva, Y. (2017). Single case studies as a prime example for exploratory research. *PsyCh Journal*, *6*(2), 107–109. https://doi.org/10.1002/pchj.176

Bao, Y., Pöppel, E., Wang, L., Lin, X., Yang, T., Avram, M., … Zaytseva, Y. (2015). Synchronization as a biological, psychological and social mechanism to create common time: A theoretical frame and a single case study. *PsyCh Journal*, *4*(4), 243–254. https://doi.org/10.1002/pchj.119

Bao, Y., Sander, T., Trahms, L., Pöppel, E., Lei, Q., & Zhou, B. (2011). The eccentricity effect of inhibition of return is resistant to practice. *Neuroscience Letters*, *500*(1), 47–51. https://doi.org/10.1016/j.neulet.2011.06.003

Bao, Y., Szymaszek, A., Wang, X., Oron, A., Pöppel, E., & Szelag, E. (2013). Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. *Cognition*, *129*(3), 579–585. https://doi.org/10.1016/j.cognition.2013.08.019

Bao, Y., von Stosch, A., Park, M., & Pöppel, E. (2017). Complementarity as generative principle: A thought pattern for aesthetic appreciations and cognitive appraisals in general. *Frontiers in Psychology*, *8*(727). https://doi.org/10.3389/fpsyg.2017.00727

Bao, Y., Wang, Z., Liang, W., Wang, Y., Pöppel, E., & Li, H. (2013). Inhibition of return at different eccentricities in the visual field share the same temporal window. *Neuroscience Letters*, *534*, 7–11. https://doi.org/10.1016/j.neulet.2012.11.046

Bao, Y., Yang, T., Lin, X., & Pöppel, E. (2016). Donders revisited: Discrete or continuous temporal processing underlying reaction time distributions? *PsyCh Journal*, *5*(3), 177–179. https://doi.org/10.1002/pchj.142

Bao, Y., Yang, T., Lin, X., Fang, Y., Wang, Y., Pöppel, E., & Lei, Q. (2016). Aesthetic preferences for eastern and western traditional visual art: Identity matters. *Frontiers in Psychology*, *7*(1596). https://doi.org/10.3389/fpsyg.2016.01596

Bao, Y., Yang, T., Zhang, J., Zhang, J., Lin, X., Paolini, M., … Silveira, S. (2017). The "third abstraction" of the Chinese artist LaoZhu: Neural and behavioral indicators of aesthetic appreciation. *PsyCh Journal*, *6*(2), 110–119. https://doi.org/10.1002/pchj.167

Barrett, K. E., Barman, S. M., Boitano, S., & Brooks, H. (2012). In NY: McGraw-Hill Medical (Ed.), *Ganong's review of medical physiology* (Vol. *24*).

Binder, J. R. (2015). The Wernicke area: Modern evidence and a reinterpretation. *Neurology*, *85*(24), 2170–2175. https://doi.org/10.1212/WNL.0000000000002219

Bleuler, M. (1969). *Lehrbuch der Psychiatrie (textbook of psychiatry)* (11th ed.). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-662-00125-7

Bodis-Wollner, I. (2008). Pre-emptive perception. *Perception*, *37* (3), 462–478. https://doi.org/10.1068/p5880

Bourguignon, M., De Tiege, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., … Jousmäki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping*, *34*(2), 314–326. https://doi.org/10.1002/hbm.21442

Chow, W. Y., Momma, S., Smith, C., Lau, E., & Phillips, C. (2016). Prediction as memory retrieval: Timing and mechanisms. *Language, Cognition and Neuroscience*, *31*(5), 617–627. https://doi.org/10.1080/23273798.2016.1160135

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, e62. https://doi.org/10.1017/s0140525x1500031x

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/s0140525x01003922

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Dick, A. S., & Andric, M. (2013). The neurobiology of receptive-expressive language interdependence. *Behavioral and Brain Sciences*, *36*(4), 352–353. https://doi.org/10.1017/s0140525x12002543

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., … Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.3389/fnhum.2017.00554

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, *81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Donhauser, P. W., & Baillet, S. (2019). Two distinct neural timescales for predictive speech processing. *Neuron*, *105*(2), 385–393. https://doi.org/10.1016/j.neuron.2019.10.019

Elbert, T., Ulrich, R., Rockstroh, B., & Lutzenberger, W. (1991). The processing of temporal intervals reflected by CNV-like brain potentials. *Psychophysiology*, *28*(6), 648–655. https://doi.org/10.1111/j.1469-8986.1991.tb01009.x

Fabb, N. (2013). There is no psychological limit on the duration of metrical lines in performance: Against Turner and Pöppel. *International Journal of Literary Linguistics*, *2*(1). https://doi.org/10.15462/ijll.v2i1.14

Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, *4*(4), 636–640. https://doi.org/10.1093/schbul/4.4.636

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, *47*(2), 250–272. https://doi.org/10.1016/s0749-596x(02)00004-9

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, *35*(1), 1–37. https://doi.org/10.1146/annurev.psych.35.1.1

Gambi, C., & Pickering, M. J. (2017). Models linking production and comprehension. In M. J. Traxler & M. Ann Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 157–182). https://doi.org/10.1002/9781118829516.ch7

Gerstner, G. E., & Cianfarani, T. (1998). Temporal dynamics of human masticatory sequences. *Physiology & Behavior*, *64*(4), 457–461. https://doi.org/10.1016/s0031-9384(98)00107-3

Gerstner, G. E., & Fazio, V. A. (1995). Evidence of a universal perceptual unit in mammals. *Ethology*, *101*(2), 89–100. https://doi.org/10.1111/j.1439-0310.1995.tb00348.x

Gerstner, G. E., & Goldberg, L. J. (1994). Evidence of a time constant associated with movement patterns in six mammalian species. *Ethology and Sociobiology*, *15*(4), 181–205. https://doi.org/10.1016/0162-3095(94)90013-2

Gómez, C., Argandoña, E. D., Solier, R. G., Angulo, J. C., & Vázquez, M. (1995). Timing and competition in networks representing ambiguous figures. *Brain and Cognition*, *29*(2), 103–114. https://doi.org/10.1006/brcg.1995.1270

Hale, J., Ward, J. A., Buccheri, F., Oliver, D., & Hamilton, A. F. D. C. (2020). Are you on my wavelength? Interpersonal coordination in dyadic conversations. *Journal of Nonverbal Behavior*, *44*(1), 63–83. https://doi.org/10.1007/s10919-019-00320-3

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Hirsh, I. J., & Sherrick, C. E. S., Jr. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, *62*(5), 423–432. https://doi.org/10.1037/e572322012-046

Kagerer, F. A., Wittmann, M., Szelag, E., & Steinbüchel, N. V. (2002). Cortical involvement in temporal reproduction: Evidence for differential roles of the hemispheres. *Neuropsychologia*, *40*(3), 357–366. https://doi.org/10.1016/s0028-3932(01)00111-7

Kien, J., & Kemp, A. (1994). Is speech temporally segmented? Comparison with temporal segmentation in behavior. *Brain and Language*, *46*(4), 662–682. https://doi.org/10.1006/brln.1994.1036

Kien, J., Schleidt, M., & Schöttner, B. (1991). Temporal segmentation in hand movements of chimpanzees (*Pan troglodytes*) and comparisons with humans. *Ethology*, *89*(4), 297–304. https://doi.org/10.1111/j.1439-0310.1991.tb00375.x

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633. https://doi.org/10.1080/01690969308407588

Kowal, S., O'Connell, D. C., & Sabin, E. J. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research*, *4*(3), 195–207. https://doi.org/10.1007/bf01066926

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Langman, V. A., Bamford, O. S., & Maloiy, G. M. O. (1982). Respiration and metabolism in the giraffe. *Respiration Physiology*, *50*(2), 141–152. https://doi.org/10.1016/0034-5687(82)90013-5

Lau, E. (2018). Neural indices of structured sentence representation: State of the art. In K. Federmeier & E. Schotter (Eds.), *Psychology of learning and motivation* (Vol. *68*, pp. 117–142): Elsevier. https://doi.org/10.1016/bs.plm.2018.08.004

Lau, E., & Liao, C. H. (2017). Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases. *Language, Cognition and Neuroscience*, *33*(5), 633–647. https://doi.org/10.1080/23273798.2017.1400081

Lei, Q., Bao, Y., Wang, B., & Gutyrchik, E. (2012). fMRI correlates of inhibition of return in Perifoveal and peripheral visual field. *Cognitive Processing*, *13*, 223–227. https://doi.org/10.1007/s10339-012-0487-3

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. https://doi.org/10.1093/acrefore/9780199384655.013.404

Marr, D. (1982). *Vision: A computational approach* (). San Francisco, CA: Freeman & Co.

Martin, B., Wittmann, M., Franck, N., Cermolacce, M., Berna, F., & Giersch, A. (2014). Temporal structure of consciousness and minimal self in schizophrenia. *Frontiers in Psychology*, *5*, 1175. https://doi.org/10.3389/fpsyg.2014.01175

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, *79*(6), 487–509. https://doi.org/10.1037/h0033467

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Meyer, L., Sun, Y., & Martin, A. E. (2019). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 1–11. https://doi.org/10.1080/23273798.2019.1693050

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–352. https://doi.org/10.1037/0033-295x.101.2.343

Miyake, Y. (2012). kyōsōteki komyunikesyon to "ma" [coordinative communication and "Ma"]. *Baiomekanisumu Gakkaishi [Journal of the Society of Biomechanisms Japan]*, *36*(2), 97–103. https://doi.org/10.3951/sobim.36.97

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, *37*(8), 2767–2783. https://doi.org/10.1002/hbm.23206

Nagy, E. (2011). Sharing the moment: The duration of embraces in humans. *Journal of Ethology*, *29*(2), 389–393. https://doi.org/10.1007/s10164-010-0260-y

Overath, T., McDermott, J. H., Zarate, J. M., & Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nature Neuroscience*, *18*(6), 903–911. https://doi.org/10.1038/nn.4021

Park, M., Gutyrchik, E., Bao, Y., Zaytseva, Y., Carl, P., Welker, L., … Meindl, T. (2014). Differences between musicians and non-musicians in neuro-affective processing of sadness and fear expressed in music. *Neuroscience Letters*, *566*, 120–124. https://doi.org/10.1016/j.neulet.2014.02.041

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.

Park, M., Gutyrchik, E., Welker, L., Carl, P., Pöppel, E., Zaytseva, Y., … Bao, Y. (2015). Sadness is unique: Neural processing of emotions in speech prosody in musicians and non-musicians. *Frontiers in Human Neuroscience*, *8*, 1049. https://doi.org/10.3389/fnhum.2014.01049

Parkes, M. J. (2006). Breath-holding and its breakpoint. *Experimental Physiology*, *91*(1), 1–15. https://doi.org/10.1113/expphysiol.2005.031625

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428. https://doi.org/10.1016/j.cogbrainres.2004.09.012

Pickering, M. J., & Clark, A. (2014). Getting ahead: Forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*(9), 451–456. https://doi.org/10.1016/j.tics.2014.05.006

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/s0140525x12001495

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, *21*(6), 1–13. https://doi.org/10.1038/s41583-020-0304-4

Polgári, P., Causin, J.-B., Weiner, L., Bertschy, G., & Giersch, A. (2020). Novel method to measure temporal windows based on eye movements during viewing of the Necker cube. *PLoS One*, *15*(1), e0227506. https://doi.org/10.1371/journal.pone.0227506

Pöppel, E. (1970). Excitability cycles in central intermittency. *Psychologische Forschung*, *34*, 1–9. https://doi.org/10.1007/BF00422860

Pöppel, E. (1972). Oscillations as possible basis for time perception. In *The study of time* (pp. 219–241). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-65387-2_16

Pöppel, E. (1988). *Mindworks: Time and conscious experience* (): Harcourt Brace Jovanovich.

Pöppel, E. (1989a). The measurement of music and the cerebral clock: A new theory. *Leonardo*, *22*, 83–89. https://doi.org/10.2307/1575145

Pöppel, E. (1989b). Taxonomy of the subjective: An evolutionary perspective. In *Neuropsychology of visual perception* (pp. 219–232). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781315441849-11

Pöppel, E. (1994). Temporal mechanisms in perception. *International Review of Neurobiology*, *37*, 185–202. https://doi.org/10.1016/s0074-7742(08)60246-9

Pöppel, E. (1997a). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, *1*(2), 56–61. https://doi.org/10.1016/s1364-6613(97)01008-5

Pöppel, E. (1997b). Consciousness versus states of being conscious. *Behavioral and Brain Sciences*, *20*(1), 155–156. https://doi.org/10.1017/s0140525x97300053

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1887–1896. https://doi.org/10.1098/rstb.2009.0015

Pöppel, E. (2010). Perceptual identity and personal self: neurobiological reflections. In T. Maruszewski, M. Fajkowska, & M. M. Eysenck (Eds.), *Personality from biological, cognitive, and social perspectives* (pp. 75–82). Clinton Corners, NY: Eliot Werner Public.

Pöppel, E., & Bao, Y. (2011). Three modes of knowledge as basis for intercultural cognition and communication: A theoretical perspective. In *Culture and neural frames of cognition and communication* (pp. 215–231). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-15423-2_14

Pöppel, E., & Bao, Y. (2012). The visual field paradox: A theoretical account on the reafference principle providing a common frame for the homogeneity and inhomogeneity of visual representation. *Cognitive Processing*, *13*(1), 285–287. https://doi.org/10.1007/s10339-012-0489-1

Pöppel, E., & Bao, Y. (2017). The power of single cases: Examples from the early visual pathway and from visual art. In *Cultural patterns and neurocognitive circuits II: East-west connections* (pp. 33–62). Singapore: World Scientific Press.

Pöppel, E., & Logothetis, N. (1986). Neuronal oscillations in the human brain. *Naturwissenschaften*, *73*(5), 267–268. https://doi.org/10.1007/bf00367781

Radilová, J., Pöppel, E., & Ilmberger, J. (1990). Auditory reversal timing. *Activitas Nervosa Superior*, *32*, 13.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Roll, M., Gosselke, S., Lindgren, M., & Horne, M. (2013). Time-driven effects on processing grammatical agreement. *Frontiers in Psychology*, *4*(1004). https://doi.org/10.3389/fpsyg.2013.01004

Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language*, *121*(3), 267–272. https://doi.org/10.1016/j.bandl.2012.03.002

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, *2*(1), 95–100. https://doi.org/10.3758/bf03197498

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience*, *5*(3), 363–370. https://doi.org/10.1162/jocn.1993.5.3.363

Schleidt, M., & Kien, J. (1997). Segmentation in behavior and what it can tell us about brain function. *Human Nature*, *8*(1), 77–111. https://doi.org/10.1007/s12110-997-1005-7

Schleidt, M., Eibl-Eibesfeldt, I., & Pöppel, E. (1987). A universal constant in temporal segmentation of human short-term behavior. *Naturwissenschaften*, *74*(6), 289–290. https://doi.org/10.1007/bf00366417

Schremm, A., Horne, M., & Roll, M. (2015). Time-driven effects on processing relative clauses. *Journal of Psycholinguistic Research*, *45*(5), 1033–1044. https://doi.org/10.1007/s10936-015-9391-1

Stefanics, G., Kremláček, J., & Czigler, I. (2014). Visual mismatch negativity: A predictive coding view. *Frontiers in*

Human Neuroscience, *8*(666). https://doi.org/10.3389/fnhum.2014.00666

Szelag, E., & Pöppel, E. (2000). Temporal perception: A key to understanding language. *Behavioral and Brain Sciences*, *23*(1), 52–52. https://doi.org/10.1017/s0140525x0055239x

Tanida, K., & Pöppel, E. (2006). A hierarchical model of operational anticipation windows in driving an automobile. *Cognitive Processing*, *7*(4), 275–287. https://doi.org/10.1007/s10339-006-0152-9

Teng, X., Ma, M., Yang, J., Blohm, S., Cai, Q., & Tian, X. (2020). Constrained structure of ancient Chinese poetry facilitates speech content grouping. *Current Biology*, *30*, 1299–1305.e7. https://doi.org/10.1016/j.cub.2020.01.059

Teuber, H. L. (1960). Perception. In *Handbook physiology neurophysiol* (Vol. *III*, pp. 1595–1668).

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*(166). https://doi.org/10.3389/fpsyg.2010.00166

Tian, X., & Poeppel, D. (2013). The effect of imagination on stimulation: The functional specificity of efference copies in speech processing. *Journal of Cognitive Neuroscience*, *25*(7), 1020–1036. https://doi.org/10.1162/jocn_a_00381

Tian, X., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex*, *77*, 1–12. https://doi.org/10.1016/j.cortex.2016.01.002

Turner, F., & Pöppel, E. (1988). Metered poetry, the brain, and time. In *Beauty and the brain* (pp. 71–90). Basel, Switzerland: Birkhäuser. https://doi.org/10.1007/978-3-0348-6350-6_4

Verduyn, P., & Lavrijsen, S. (2015). Which emotions last longest and why: The role of event importance and rumination. *Motivation and Emotion*, *39*(1), 119–127. https://doi.org/10.1007/s11031-014-9445-y

Vollrath, M., Kazenwadel, J., & Krüger, H. P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften*, *79*(10), 479–480. https://doi.org/10.1007/bf01139205

von Helmholtz, H. (1896). *Handbuch der physiologischen Optik (handbook of physiological optics)* (). Hamburg und Leipzig: Verlag von Leopold Voss, 2 Auflage.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystern und Peripherie (the Reafference principle. Interactions between the central nervous system and the periphery). *Naturwissenschaften*, *37*, 464–476. https://doi.org/10.1007/bf00622503

von Holst, E., & Mittelstaedt, H. (1971). The principle of reafference: Interactions between the central nervous system and the peripheral organs. In P. Dodwell (Ed.), *Perceptual processing: Stimulus equivalence and pattern recognition* (pp. 41–72).

von Steinbüchel, N. (1998). Temporal ranges of central nervous processing: Clinical evidence. *Experimental Brain Research*, *123*(1–2), 220–233. https://doi.org/10.1007/s002210050564

Wagers, M. W., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, *45*(2), 395–433. https://doi.org/10.1017/s0022226709005726

Wang, L., Bao, Y., Zhang, J., Lin, X., Yang, L., Pöppel, E., & Zhou, B. (2016). Scanning the world in three seconds: Mismatch negativity as an indicator of temporal segmentation. *PsyCh Journal*, *5*(3), 170–176. https://doi.org/10.1002/pchj.144

Wang, L., Kuperberg, G., & Jensen, O. (2018). Specific lexicosemantic predictions are associated with unique spatial and temporal patterns of neural activity. *eLife*, *7*, e39061. https://doi.org/10.1080/01690969308407588

Wang, L., Lin, X., Zhou, B., Pöppel, E., & Bao, Y. (2015). Subjective present: A window of temporal integration indexed by mismatch negativity. *Cognitive Processing*, *16*(1), 131–135. https://doi.org/10.1007/s10339-015-0687-8

Wang, M., Arteaga, D., & He, B. J. (2013). Brain mechanisms for simple perception and bistable perception. *Proceedings of the National Academy of Sciences*, *110*(35), E3350–E3359. https://doi.org/10.1073/pnas.1221945110

White, P. A. (2017). The three-second "subjective present": A critical review and a new proposal. *Psychological Bulletin*, *143* (7), 735–756. https://doi.org/10.1037/bul0000104

Wittmann, M. (1999). Time perception and temporal processing levels of the brain. *Chronobiology International*, *16*(1), 17–32. https://doi.org/10.3109/07420529908998709

Wurtz, R. H. (2018). Corollary discharge contributions to perceptual continuity across saccades. *Annual Review of Vision Science*, *4*, 215–237. https://doi.org/10.1146/annurev-vision-102016-061207

Zaytseva, Y., & Bao, Y. (2015). Special collection on single case studies. *PsyCh Journal*, *4*(4), 177–177. https://doi.org/10.1002/pchj.122

Zhang, Y., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 765–781. https://doi.org/10.1037/a0019254

Zhao, C., Zhang, D., & Bao, Y. (2018). A time window of 3 s in the aesthetic appreciation of poems. *PsyCh Journal*, *7*(1), 51–52. https://doi.org/10.1002/pchj.194

Zhou, B., Bao, Y., Sander, T., Trahms, L., & Pöppel, E. (2010). Dissociation of summation and peak latencies in visual processing: An MEG study on stimulus eccentricity. *Neuroscience Letters*, *483*(2), 101–104. https://doi.org/10.1016/j.neulet.2010.07.070

Zhou, B., Pöppel, E., & Bao, Y. (2014). In the jungle of time: The concept of identity as a way out. *Frontiers in Psychology*, *5*(844). https://doi.org/10.3389/fpsyg.2014.00844

Zhou, B., Pöppel, E., Wang, L., Yang, T., Zaytseva, Y., & Bao, Y. (2016). Seeing without knowing: Operational principles along the early visual pathway. *PsyCh Journal*, *5*(3), 145–160. https://doi.org/10.1002/pchj.141

Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). Interpretations of frequency domain analyses of neural entrainment: Periodicity, fundamental frequency, and harmonics. *Frontiers in Human Neuroscience*, *10*(274). https://doi.org/10.3389/fnhum.2016.00274

© 2020 The Authors. PsyCh Journal published by Institute of Psychology, Chinese Academy of Sciences and John Wiley & Sons Australia, Ltd.