

Visual apparent motion can be modulated by task-irrelevant lexical information

Lihan Chen · Xiaolin Zhou

Published online: 15 January 2011
© Psychonomic Society, Inc. 2011

Abstract Previous studies have repeatedly demonstrated the impact of Gestalt structural grouping principles upon the parsing of motion correspondence in ambiguous apparent motion. Here, by embedding Chinese characters in a visual Ternus display that comprised two stimulus frames, we showed that the perception of visual apparent motion can be modulated by activation of task-irrelevant lexical representations. Each frame had two disks, with the second disk of the first frame and the first disk of the second frame being presented at the same location. Observers could perceive either “element motion,” in which the endmost disk is seen as moving back and forth while the middle disk at the central position remains stationary, or “group motion,” in which both disks appear to move laterally as a whole. More reports of group motion, as opposed to element motion, were obtained when the embedded characters formed two-character compound words than when they formed nonwords, although this lexicality effect appeared to be attenuated by the use of the same characters at the overlapping position across the two frames. Thus, grouping of visual elements in a changing world can be guided by both structural principles and prior world knowledge, including lexical information.

Keywords Visual apparent motion · Ternus display · Perceptual grouping · Top-down modulation · Lexical representation

L. Chen
Center for Brain and Cognitive Sciences
and Department of Psychology, Peking University,
Beijing, China

X. Zhou (✉)
Center for Brain and Cognitive Sciences and Department
of Psychology, and Key Laboratory of Machine Perception
and Key Laboratory of Computational Linguistics
(Ministry of Education), Peking University,
Beijing 100871, China
e-mail: xz104@pku.edu.cn

Apparent motion is a *perceived* motion produced by the presentation of sequential but static images containing no physically continuous motion. Two components of motion parsing involved in apparent motion are object identity and motion correspondence, with the former referring to what the object is and the latter indicating the spatiotemporal relationship between objects presented consecutively (Kolers, 1972; Ternus, 1926; Ullman, 1979; Wertheimer, 1912). Previous studies have identified two categories of factors that influence motion correspondence (Tse, Cavanagh, & Nakayama, 1998): structural factors (i.e., factors related to bottom-up grouping) and semantic factors (i.e., factors related to top-down processing). A great number of studies focused on the former category, where the motion correspondence can be affected by spatial proximity (Burt & Sperling, 1981; Kolers, 1972; Kramer & Yantis, 1997; Shechter, Hochstein, & Hillman, 1988; Ullman, 1979), stimulus orientation (Green, 1986), luminance contrast (Alais & Lorenceau, 2002; He & Ooi, 1999; Ma-Wyatt, Clifford, & Wenderoth, 2005), topological characteristics (Chen, 1985; Prazdny, 1986), temporal proximity, or temporal cuing from other sensory modalities (Chen, Shi, & Müller, 2010; Kramer & Yantis, 1997; Shi, Chen, & Müller, 2010). In contrast, few studies have investigated the influence of semantic factors on motion correspondence. Tse and Cavanagh (2000) showed that orthographic knowledge with Chinese characters (e.g., 当, /dang/) can affect the line motion illusion. Although each stroke in a character was presented all at once, for the last stroke at the bottom of the character (e.g., “-”), observers raised in China tended to see apparent motion in the direction following the writing sequence of the stroke (i.e., from left to right), while American observers not familiar with Chinese characters tended to see line motion in the opposite direction, consistent with the Gestalt principle of continuation. Thus, orthographic knowledge or top-down expectation is capable of overriding bottom-up cues to apparent motion.

The Ternus display provides a fascinating demonstration of motion correspondence between two visual frames (Petersik & Rice, 2006; Ternus, 1926) and an opportunity to investigate top-down influence on visual apparent motion (Yu, 2000). The Ternus display is an ambiguous display in which multiple correspondences are possible (Fig. 1a and b). In the present setup, the first frame was composed of two dots presented at Locations 1 and 2, followed by the second frame, presented at Locations 2 and 3. Observers could perceive either “element motion,” in which the endmost dot is seen as moving back and forth while the middle dot at the overlapping or central position remains stationary (flashing), or “group motion,” in which both dots appear to move in a manner of lateral displacement as a whole (Ternus, 1926). It has been proposed that in Ternus apparent motion, temporal and spatial grouping processes are in competition (He & Ooi, 1999; Kramer & Yantis, 1997; Petersik & Rice, 2006). At a short interstimulus interval (ISI) between the two frames, temporal grouping (temporal proximity) prevails; that is, the stimulus at the “overlapping” location (i.e., the second dot) of the first frame is likely to be grouped with the stimulus appearing at the same location (the first dot) in the second frame, leading to the percept of element motion. In contrast, at a long ISI, temporal proximity weakens, and spatial grouping within each frame becomes more prominent, giving rise to a dominant percept of group motion. However, evidence has also shown that in Ternus displays with an ISI of 0 ms, certain structural relationships between the elements of the first and second frames can be manipulated to give either element or group motion in the

absence of any change in temporal grouping (Scott-Samuel & Hess, 2001; Wallace & Scott-Samuel, 2007).

By presenting observers with a cartoon of a walking person’s feet (two circles) or of the wheels of a car, Yu (2000) examined whether such semantic context could affect the proportion of Ternus displays perceived as group or element motion. The results showed that there were more reports of element motion after seeing a walking person’s feet than after seeing the wheels of a car. The author concluded that the observers’ knowledge of human walking with one foot remaining static and the other moving and of wheels of a car moving as one unit affected the parsing of motion correspondence. The aim of the present study was to continue this line of research by showing that task-irrelevant lexical information can affect the perception of a Ternus display. In Yu’s study, world knowledge priming was coupled with structural consistency between the prime (a cartoon of a walking person or a car) and the target (the remaining feet or wheels of the partially occluded picture). Thus, the contextual or semantic effect on the parsing of motion correspondence was based on this structural invariance between the prime and the Ternus display. In the present study, we took advantage of the characteristics of Chinese compound words (words composed of two or more words or morphemes, such as *bathroom*, *butterfly*, *bookstore*, etc.) and manipulated the lexicality of compounds whose constituent morphemes (characters) were embedded in the visual disks of the Ternus display (Fig. 1c–h). Most Chinese lexical items are two-character compound words, with a character standing for a morpheme and corresponding to a syllable in spoken form (Zhou & Marslen-Wilson, 2009;

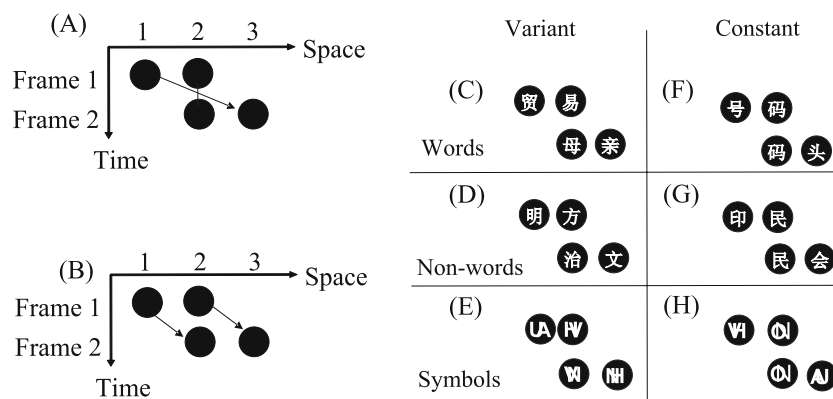


Fig. 1 Ternus apparent motion paradigm and exemplar stimuli. (a) “Element motion” percept: The middle dot is perceived as staying at the same location (“2”) across the two frames, while the outer dot is perceived as moving from one location to the other (from “1” to “3”). (b) “Group motion” percept: The two dots are perceived as moving together in the manner of a coherent lateral displacement. (c) Variant–word display: Disks were embedded with two two-character compound words (“trade” and “mother,” in this example). (d) Variant–nonword display: The embedded characters could not form meaningful com-

pounds. (e) Variant–symbol display. (f) Constant–word display: The second character in the first frame and the first character in the second frame were the same, whereas the two compound words were not semantically related (“number” and “dock,” in this example). (g) Constant–nonword display: The second character in the first frame and the first character in the second frame were the same; both pairs of characters could not form meaningful compound words. (h) Constant–symbol display: The middle symbol remained the same across the two frames

Zhou, Marslen-Wilson, Taft, & Shu, 1999). Each character is written with a finite set of basic strokes (e.g., horizontal, vertical, left-falling, dot, and turning strokes) that are combined in a systematic way. Previous studies demonstrated that two-character Chinese compound words are represented as wholes in the lexicon (e.g., Zhou et al., 1999). A recent study showed that this whole-word representation can be used as top-down guidance in attentional selection by grouping constituent morphemes into integrated objects (Li & Logan, 2008). Using a variant of the two-rectangle cuing paradigm (Egly, Driver, & Rafal, 1994), Li and Logan found that the color on a target character was detected faster if this character and the cuing character formed a compound word than if they did not. For this study, if the perception of a Ternus display can be affected by top-down guidance of lexical information, we would expect to see more reports of group motion when the two characters in each frame form a meaningful compound word (an integrated object) than when they do not. Moreover, this lexicality effect could be attenuated by the constant display in the middle element, with the same character at the middle location across the two frames (Fig. 1f and g). Previous studies have shown that a constant display in the central disk can lead to fewer reports of group motion (Braddick & Adlard, 1978; Dawson, Nevin-Meadows, & Wright, 1994; Kramer & Rudd, 1999; Petersik, Hicks, & Pantle, 1978; Petersik & Rice, 2008). This phenomenon was interpreted as due to stronger temporal (structural) grouping between the same elements at the central position across the two frames. With this setting, we pitted the high-level lexical representation against the low-level structural grouping to reveal how the two factors might interact in parsing motion correspondence.

Method

Participants

Twenty-four undergraduate or graduate students (14 females) took part in the experiment. They were right-handed, had normal or corrected-to-normal vision, and had no known cognitive or neurological disorders. Their ages ranged from 18 to 32 years. Informed consent was obtained from each participant as required by the Academic Affairs Committee, Department of Psychology at Peking University.

Design and materials

Each visual display was composed of two stimulus frames, with each frame composed of two black disks. The critical stimuli were embedded in the disks, which were presented on a gray background. A 2 (display type: variant and constant) \times 3 (stimulus category: words, nonwords, and

symbols) \times 7 (ISI between frames: 50, 80, 110, 140, 170, 200, and 230 ms) factorial design was used. The Ternus display had two types: “variant,” where the stimulus embedded in the middle disk changed across the two frames, and “constant,” where the stimuli embedded in the middle disk remained to be the same across the two frames.

The embedded stimuli in each frame consisted of (a) two-character compound words (e.g., *bathroom*), (b) compound nonwords composed of randomly combined characters (e.g., *bath-fly*), or (c) nonsense stimuli composed of combined letters. Each compound word or nonsense word contained two meaningful Chinese characters, with each character composed of several strokes. For the word stimuli, we selected seven groups of compounds, each corresponding to one ISI condition. Each group was divided further into two subgroups (with 24 pairs of words each) corresponding, respectively, to the constant and the variant displays. The average word frequencies for the constant and variant displays, drawn from Cai and Brysbaert (2010), were 0.4206 and 0.4449 per million. An ANOVA with display type and order of frame as two factors found no main effect of either display type, $F(1, 668) = 0.18$, $p > .1$, or order of frame, $F(1, 668) = 0.13$, $p > .1$, and no interaction between display type and frame, $F(1, 668) = 0.33$, $p > .1$. The average number of strokes, which represents the visual complexity of the compounds, was also matched for the two types of displays and for the two frames, with no effect of either the display type (14.87 and 14.96 per word, respectively, for the variant and constant displays), $F(1, 668) = 0.14$, $p > .1$, or frame (14.85 and 14.97 per word, respectively, for the first and second frames), $F(1, 668) = 0.30$, $p > .1$, and no interaction between display type and frame, $F(1, 668) = 1.47$, $p > .1$.

Words in the two frames were not related semantically, phonologically, or orthographically. The semantic relatedness between the paired words was checked postexperiment by asking the 24 participants to rate each pair on a 7-point Likert scale (1 = *not related at all*, 7 = *highly related*). The average scores were 2.54 and 2.37 for word pairs used, respectively, for the constant and variant displays, $t(23) = 1.468$, $p = .156$.

Seven groups of nonwords were created by pseudorandomly recombining the characters used in the compound words for each ISI group. In this way, the same set of characters were used for word and nonword stimuli. The nonsense symbols were created by randomly overlapping two Roman letters (drawn from H, V, N, O, U, A, with a size of 2.4° of visual angle for the combined symbols).

Each black disk (3.15 cd/m² in luminance) subtended 3.2° in visual angle, and the distance between the centers of the two adjacent disks was 4.5° in visual angle. The embedded characters were in bold and white, with a size of 2.4° of visual angle. The luminance of the screen background was 9.30 cd/m². The average luminance of

the visual disks with characters was 13.2 cd/m². The two frames shared one element location (the middle location) at the center of the monitor but contained two other elements located at horizontally opposite positions relative to the center. The motion direction of the Ternus frames was always from left to right.

Visual stimuli were presented on a 22-in. CRT monitor with a refresh rate of 100 Hz and a resolution of 1,024 × 768 pixels. The presentation of stimuli was controlled by a computer program that was developed with MATLAB (MathWorks Inc., Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and was run on a computer with the Intel Core 2 Quad Q8300 processor and an Nvidia GeForce 9600 GT graphics card. Participants were tested individually in a soundproof, dimly lit chamber, with the viewing distance kept at 57 cm and maintained by the use of a chinrest. The ambient luminance in the chamber was 0.05 cd/m².

Procedure

The 42 subgroups of stimuli (2 × 3 × 7) were randomly combined for each participant. The 1,008 trials (42 × 24) were put into a random sequence and divided into 12 blocks, with each block containing 84 trials.

Each trial began with a central fixation cross presented for 300 to 500 ms. The purpose of using a variable presentation time was to reduce time-based expectations toward the next stimulus. The two visual frames in each Ternus display were then presented, with each frame presented for 50 ms and with the ISI between frames varying from 50 to 230 ms. A blank screen was then presented for 300 to 500 ms, followed by a question mark that prompted the participant to make a two-alternative forced choice (2AFC) response with the mouse button, indicating whether he or she had perceived element or group motion. About 500 to 700 ms after the response, the next trial began. A short break was given between blocks.

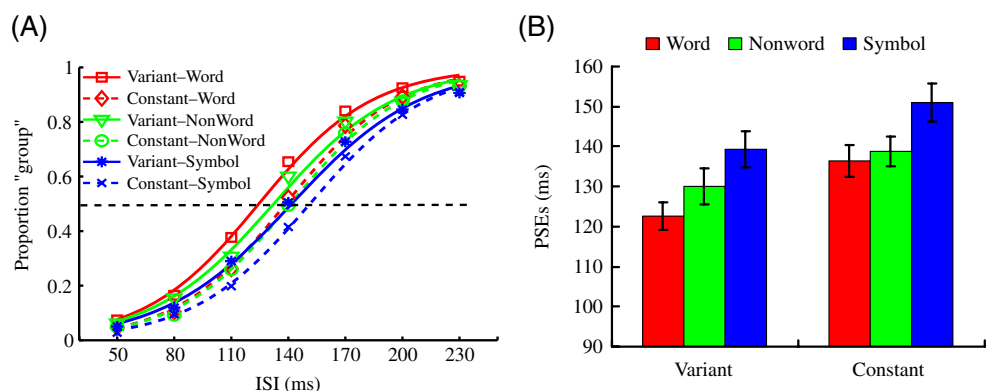
Before the formal test, the participants received a practice block of stimuli containing frames composed of only black disks (i.e., without characters or symbols). In the practice session, the Ternus display (without embedded stimuli) had two levels of ISI, one being 50 ms and the other 260 ms, with the duration of each frame being 50 ms. For the ISI of 50 ms, the display was obviously in “element motion”; for the ISI of 260 ms, the display was obviously in “group motion.” If the participant made an incorrect response, feedback was displayed on the screen. The practice session went on until the accuracy of response was above 90%. Nearly all of the participants attained this accuracy after about 50 trials. The purpose of giving feedback was to make sure that the participant was familiar with the categorization task and with the experimental procedure. The participant then received a further practice block of 84 trials, which contained stimuli similar to the formal test and in which no feedback was given.

Results

For each experimental condition, the percentage of “group motion” responses was calculated individually for each participant. For each combination between display and stimulus type, the seven data points (one for each ISI) were fitted into the psychometric curve using a logistic function (Treutwein & Strasburger, 1999). For each participant, the ISI at which he/she was equally likely to report the two different motion percepts was calculated by estimating the point of subjective equality (PSE) from the psychometric curve. Across ISIs, the lower the PSE, the more likely the display was to be perceived as group motion. The average psychometric curves from all of the participants are depicted in Fig. 2a, and the calculated PSEs, on which the statistical analysis was based, are presented in Fig. 2b.

An ANOVA with display type and stimulus category as two within-participant factors showed a significant main

Fig. 2 (a) Distribution of proportions of “group motion” responses as a function of stimulus category, display type, and ISI. (b) Transition thresholds (PSEs) for “group motion” responses. The PSEs correspond to the ISI values acquired from the intersection points with the .5 line in panel A



effect of display type, $F(1, 23) = 34.44$, $p < .001$, $\eta_p^2 = .600$, with the PSE lower for constant displays (130.7 ms) than for variant displays (142.0 ms). More importantly, the main effect of stimulus category was significant, $F(2, 46) = 17.786$, $p < .001$, $\eta_p^2 = 0.436$, although the interaction between stimulus category and display type was not, $F(2, 46) = 1.143$, $p = .328$, $\eta_p^2 = .047$. Pairwise comparisons, with Bonferroni correction, were conducted. Both word (129.5 ms) and nonword (134.5 ms) stimuli had lower PSEs than symbol stimuli (145.1 ms), $t(23) = 5.025$, $p < .001$, and $t(23) = 3.605$, $p < .01$, respectively. Crucially, the difference between word and nonword stimuli was also significant, $t(23) = 2.833$, $p < .05$. Further comparisons showed that, for the constant displays, the PSE was smaller for word stimuli (122.6 ms) than for nonword stimuli (130.1 ms), $t(23) = 2.91$, $p < .01$; for the variant displays, the difference between the two types of stimuli (136.3 vs. 138.9 ms) was not significant, $t(23) = 1.358$, $p = .188$. Therefore, Ternus displays with meaningful characters were more likely to be perceived as being in group motion, relative to those with meaningless symbols. Moreover, this report of group motion was more likely to occur for compound words than for compound nonwords, although it appeared that this lexicality effect was modulated by display type.

Discussion

By embedding task-irrelevant compound words and nonwords in the disks of a Ternus display, this experiment demonstrated that activation of the whole-word representation of a compound word can affect the perception of visual apparent motion, even though this lexical processing is completely task-irrelevant. This finding, while replicating and extending those of Yu (2000), provides strong evidence for a top-down semantic impact on the motion correspondence in illusory perception.

Compared with symbol displays, nonword displays had lower PSEs. There could be two accounts for this differential effect. One account attributes it to the difference in separating the two frames. Studies have indicated that, as compared with Ternus disks with identical visual texture features, Ternus displays that shared lower similarity at the middle location across two frames had more reports of group motion (Braddick & Adlard 1978; Petersik et al., 1978; Petersik & Rice, 2008; Scott-Samuel & Hess, 2001). It is clear from Fig. 1 that, generally, characters used in the two frames had lower visual similarity than did nonsense symbols embedded in the two frames. Moreover, activation of the semantic representations of characters (morphemes), which were clearly different for different characters or frames, could provide an additional cue to distinguish the frames. Thus, the dissimilarity between the two frames for

compound nonwords led to lower PSEs than for symbols. Alternatively, because our compound nonwords were composed of meaningful morphemes or characters, the processing of individual characters may have attracted more attentional resources to the frames than did the processing of the symbols. A previous study demonstrated that having deeper processing of the visual frames leads to more reports of group motion (Petersik & Pantle, 1979).

Importantly, we found a lexicality effect in the reports of group motion. Frames embedded with two-character compound words were, in general, more likely to be perceived as being in group motion than were frames embedded with randomly combined characters. We believe that this lexicality effect comes from top-down semantic guidance, since compound words are represented as wholes in the lexicon (Zhou et al., 1999) and these whole-word representations can be used to group characters into integrated objects (Li & Logan, 2008) and to guide perceptual grouping of visual objects. The two disks in each frame are more likely to be perceived as being bound together if the embedded characters form a compound word than if they form a nonword. This semantic binding leads to the percept of group motion, even though the semantic activation of the compound words is completely task-irrelevant.

Although we did not find statistical significance for the interaction between stimulus category and display type, it appears in Fig. 2b and in separate analyses that the lexicality effect was reduced in constant displays. Previous studies showed that, as compared with having different elements at the middle location over the two frames, having the same element increases cross-frame temporal grouping, leading to fewer reports of group motion (Braddick & Adlard, 1978; Dawson et al., 1994; Kramer & Rudd, 1999; Petersik et al., 1978; Petersik & Rice, 2008; Scott-Samuel & Hess, 2001). When we pitted top-down guidance of semantic activation against this bottom-up spatiotemporal Gestalt grouping, as illustrated in Fig. 1f, it seemed that the structural grouping had the upper hand in determining the percept of visual apparent motion.

It should be noted that, compared with other studies (e.g., Kramer & Yantis, 1997; Petersik & Rice, 2008), the transient thresholds (PSEs) in this study were relatively high and the effect size was relatively small. We speculate that this difference is due to the difference in frame duration, which was usually fixed at 200 ms in previous studies but was shorter in the present study (50 ms). Longer frame durations and deeper processing may lead to lower PSEs (Petersik & Pantle, 1979).

To conclude, by embedding Chinese two-character compound words in a visual Ternus display, we demonstrated that the parsing of ambiguous visual apparent motion can be modulated by the activation of task-irrelevant lexical representation. Grouping of visual elements in the changing world

around us can be guided not only by Gestalt structural principles, but also by our world knowledge, including lexical representations.

Author Note This study was supported by grants from the Natural Science Foundation of China (30770712, 30970889, 90920012), the Ministry of Science and Technology of China (2010CB833904), and the China Postdoctoral Science Foundation (20100470). We thank four anonymous reviewers for their constructive comments and helpful suggestions.

References

- Alais, D., & Lorenceau, J. (2002). Perceptual grouping in the Ternus display: Evidence for an “association field” in apparent motion. *Vision Research*, *42*, 1005–1016.
- Braddick, O. J., & Adlard, A. (1978). Apparent motion and the motion detector. In J. Armington, J. Krauskopf, & B. R. Wooten (Eds.), *Visual psychophysics and physiology* (pp. 417–426). New York: Academic Press.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, *88*, 171–195.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*, e10729.
- Chen, L. (1985). Topological structure in the perception of apparent motion. *Perception*, *14*, 197–208.
- Chen, L., Shi, Z., & Müller, H. J. (2010). Influences of intra- and crossmodal grouping on visual and tactile Ternus apparent motion. *Brain Research*, *1354*, 152–162.
- Dawson, M. R., Nevin-Meadows, N., & Wright, R. D. (1994). Polarity matching in the Ternus configuration. *Vision Research*, *34*, 3347–3359.
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161–177.
- Green, M. (1986). What determines correspondence strength in apparent motion. *Vision Research*, *26*, 599–607.
- He, Z. J., & Ooi, T. L. (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, *28*, 877–892.
- Kolers, P. A. (1972). *Aspects of motion perception*. Oxford: Pergamon Press.
- Kramer, P., & Rudd, M. (1999). Visible persistence and form correspondence in Ternus apparent motion. *Perception & Psychophysics*, *61*, 952–962.
- Kramer, P., & Yantis, S. (1997). Perceptual grouping in space and time: Evidence from the Ternus display. *Perception & Psychophysics*, *59*, 87–99.
- Li, X., & Logan, G. D. (2008). Object-based attention in Chinese readers of Chinese words: Beyond Gestalt principles. *Psychonomic Bulletin & Review*, *15*, 945–949.
- Ma-Wyatt, A., Clifford, C. W., & Wenderoth, P. (2005). Contrast configuration influences grouping in apparent motion. *Perception*, *34*, 669–685.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Petersik, J. T., Hicks, K. L., & Pantle, A. J. (1978). Apparent movement of successively generated subjective figures. *Perception*, *7*, 371–383.
- Petersik, J. T., & Pantle, A. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Research*, *19*, 143–154.
- Petersik, J. T., & Rice, C. M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*, 807–821.
- Petersik, J. T., & Rice, C. M. (2008). Spatial correspondence and relation correspondence: Grouping factors that influence perception of the Ternus display. *Perception*, *37*, 725–739.
- Prazdny, K. (1986). What variables control (long-range) apparent motion. *Perception*, *15*, 37–40.
- Shechter, S., Hochstein, S., & Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Research*, *28*, 1013–1021.
- Scott-Samuel, N. E., & Hess, R. F. (2001). What does the Ternus display tell us about motion processing in human vision? *Perception*, *30*, 1179–1188.
- Shi, Z., Chen, L., & Müller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: The influence of time interval. *Experimental Brain Research*, *203*, 723–735.
- Ternus, J. (1926). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, *7*, 81–136.
- Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics*, *61*, 87–106.
- Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition*, *74*, B27–B32.
- Tse, P. U., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. In T. Watanabe (Ed.), *High-level motion processing-computational, neurobiological and psychophysical perspectives* (pp. 249–266). Cambridge, MA: MIT Press.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Wallace, J. M., & Scott-Samuel, N. E. (2007). Spatial versus temporal grouping in a modified Ternus display. *Vision Research*, *47*, 2353–2366.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, *61*, 161–265.
- Yu, K. (2000). Can semantic knowledge influence motion correspondence? *Perception*, *29*, 693–707.
- Zhou, X., & Marslen-Wilson, W. D. (2009). Pseudohomophone effects in processing Chinese compound words. *Language and Cognitive Processes*, *24*, 1009–1038.
- Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese. *Language and Cognitive Processes*, *14*, 525–565.