An Eye-tracking Technique—Visual-World Paradigm

Guiyun Guan
Northwest Minzu University
Lanzhou, Gansu 730030, China

Axu Hu[1], Lei Guo[1], Hongzhi Yu[1]
[1]Key Laboratory of China's Ethnic Languages and
Information Technology of Ministry of Education,
Northwest Minzu University
Lanzhou, Gansu 730030, China

Abstract—The visual-world Paradigm was first developed by Cooper in 1974. Then the paradigm was further developed by Tanenhaus in 1995 by putting it into the study of syntax disambiguation in spoken language understanding. Because the paradigm is easy to carry on and doesn't have special requirements on the participants, that is, illiterate people, young children, elderly people and even some kinds of patients can participate in the experiments, the paradigm has been used widely in the study of spoken language, as well as word recognition and syntax understanding. The paper introduces the development of visual-world paradigm and its application to word recognition and syntactic processing in spoken language understanding experiments, and then analyzes its advantages and challenges, aiming to explain the eye-tracking technique — visual-world paradigm.

Keywords—eye tracking technique; visual-world paradigm; syntactic processing; advantages; limitation

I. INTRODUCTION

Visual-world Paradigm, as a very useful eye-tracking technique, means "participants listen to a spoken utterance and simultaneously look at a visual scene containing various objects while their eye movements are monitored."(Berends, et all., 2015) It is now often used in the experiments of spoken language and describing the concurrent visual world. The participants are asked to listen to a word or a sentence or a paragraph, meanwhile in front of them are some pictures with objects on. At the same time, the eye tracker is used to track their eye movements. The researchers can take advantage of the data collected by the eye tracker to deduce the participants' inner psychology of perception. The paradigm's theoretical hypotheses lies in the fact that eye movements can reflect one's process of understanding; if a word is activated, the fixation on the object related to the word will be more than that on any other objects. The use of visual-world paradigm provides a new and more effective way of studying word recognition and syntactic processing in spoken language understanding.

II. THE DEVELOPMENT OF VISUAL-WORLD PARADIGM

It was first established by Cooper in 1974. In his experiment, Cooper asked the participants to listen to a short story and told them that their pupil size would be recorded, but they could look at whatever they wanted. The results of the experiment showed that the participants' eye movements were closely "time-locked to the text"(Huettig, et al., 2010). That is, they paid more attention to the objects they heard in the story or those related to the objects they heard. Cooper called this "visual-world paradigm". But at that time the paradigm didn't attract much attention and was ignored by researchers for more than twenty years. Not until 1995 when Tanenhaus, Spivey-Knowlton, Eberhard and Sedivy further developed the paradigm, putting the paradigm into the study of how visual context influenced auditory word recognition and mediated syntactic processing did the psycholinguists begin to focus on the relationship between eye movements and spoken language processing. In China, Qiu wrote an article (Qiu, et al., 2009) about the visual-world paradigm in spoken language understanding. Then more and more Chinese researchers began to use the paradigm in their experiments.

III. THE APPLICATION OF VISUAL-WORLD PARADIGM TO WORD RECOGNITION

Word recognition, also known as lexical access, refers to "the ability of a reader to recognize written words correctly and virtually effortlessly". Generally speaking, word recognition can be divided into spoken word recognition and visual word recognition.Visual-world paradigm is a very useful technique used in the experiments of word recognition.

A. Activation of Phonological Representation

In any language understanding, word recognition is the first and most important step. There are different opinions on which part of a word can be firstly activated. Marslen-Wilson first developed Cohort Model in 1980s (Marslen-Wilson, 1978, 1980, 1984, 1987) to emphasize the "left-to-right" nature of speech recognition and study how the process of recognition unfolds as time goes on. Cohort Model means "a word is recognized via a successive reduction in the number of possible word candidates as each new phoneme is perceived." (Taft, et al., 1986) McClelland and Elman developed TRACE model in 1986 (Elman & McClelland, 1986; McClelland & Elman, 1986), giving a computational implementation of some of the ideas first expressed in the Cohort model. Norris (Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997) developed the Shortlist Model which makes use of a competition mechanism to perform bottom-up lexical segmentation.

In 1998, Allopenna and his partners first put the visual-world paradigm into the experiment of word recognition. In their experiments, participants were asked to listen to a sentence, for example "Pick up the beaker; now put it below the diamond" and then look at pictures of four objects: a target

object (e.g., beetle), a rhyme competitor (e.g., speaker), and an unrelated competitor (e.g., carriage). The results proved the existence of cohort effect and rhyme effect in the process of word recognition: the cohorts and rhymes compete for lexical activation.

In 1999, Altmann and Kamide used visual-world paradigm to study the prediction in language comprehension. The participants heard the sentence, for example, "The boy will eat..." and meanwhile saw a scene with a CAKE and other inedible objects. The results proved the prediction in spoken language comprehension.

In 2007, McQueen and Viebahn developed a version of the visual-world paradigm: using printed words to take the place of pictures, which makes it easier to design the experimental materials and widens the controlled sets of materials (only concrete noun before).

B. Activation of Semantic Representation

A word consists of form(including sound and spelling) and semantic meaning. Now that phonological representation has been proved activated in word recognition under the visual-world paradigm, then how about the semantic representation?

Huettig and his partners (2005) asked the participants to watch three groups of pictures respectively (a piano was depicted amongst unrelated distractors; a trumpet was depicted amongst those same distractors; both the piano and trumpet were depicted), meanwhile they heard the word "piano". Under the first two circumstances, the fixations to the piano and the trumpet rose, but under the last one, only fixations to the piano rose. It proved the activation of semantic representation.

Shen, Qu and Li (2016) did experiments on Chinese participants, finding that written materials simultaneously activated both meaning and sound, and the activation of meaning was stronger that that of sound.

There are still many other experiments, (Yee, et al., 2006; Liu, 2018) drawing the same conclusion of the activation of semantic representation.

C. Activation of Shape

Dahan and Tanenhaus (2004b) found that when the participants heard the referent word, e.g., "snake", they were more likely to fixate the visual competitor (e.g., rope) than to fixate either of the distractor objects (e.g., a couch and an umbrella), which proved the activation of shape during the word recognition.

Rommers and his partners (2013) did experiments, finding that the participants tended to fixate both the target (e.g., moon) and the shape competitor (e.g., tomato) more often than they fixated the unrelated control objects, which proved that the participants had anticipatorily activated the shape of the word's referent.

IV. THE APPLICATION OF VISUAL-WORLD PARADIGM TO SYNTACTIC PROCESSING

Syntactic processing has always been researchers' focus in recent decades. All the researchers do agree that visual context plays an important part in syntax processing, but they are unable to decide during which period (the beginning period, or the later period) of language processing visual context takes effect.

Tanenhaus and his partners did experiments, finding that visual context affected spoken word recognition and mediated syntactic processing, even during the earliest moments of language processing.(Tanenhaus, et al., 1995)

Chambers, Tanenhaus and Magnuson did two experiments, making use of constraint-based approaches (here the constraint referred to the nonlinguistic constraints: plausibility and referential context), to "investigate how perceptual and action-based knowledge is used in the course of language understanding and to evaluate whether this information constrains the earliest moments of syntactic processing."(Chambers, et al., 2004) The results showed that any common sense related to actions could influence syntactic processing at any time.

There are still other studies on syntactic processing by using the visual-world paradigm. (Snedeker & Trueswel, 2004; Novick, et al., 2008, Contemori, et al., 2018)

V. ADVANTAGES AND LIMITATIONS OF VISUAL- WORLD PARADIGM

The visual-world paradigm combines the visual and listening stimuli together, making it the most effective method in the experiments of spoken and listening language understanding. It has many advantages.

First, it is an online method. With the help of eye tracker, the paradigm can reveal the listeners' word processing in spoken language comprehension when they are completely unaware of it. It provides us with the proof of people's word processing mechanism. Second, the paradigm has very low requirements on the participants' knowledge background, understanding capacity and operation ability, so the preschool children, elderly people, and even some patients can be the participants. Third, the participants don't have any tasks during the visual-world paradigm experiments so they don't use strategies, which makes the experimental results closer to real word processing. Fourth, the data collected are very accurate, providing precision in determining the response time.

Of course, just like a sword with two edges, visual-world paradigm also has its own limitations.

According to Dahan (2004b), the serious limitation of visual-world paradigm is that it is a "closed set", which means "it requires use of a circumscribed visual world that is most often perceptually available to participants before hearing spoken input." So the lexical candidates in the experiments are limited. In addition, the paradigm combines listening and visual world, so the results don't only indicate the process of spoken language understanding, but also the influences of visual world.

VI. Conclusion

Eye movements can reflect one's mental processes accompanying language comprehension, and eye tracker just takes advantage of the point to catch minute eye movements. Visual-world paradigm is a very useful eye-tracking technique used by many psychologists as well as linguists to explore the inner mechanism in the process of perception. Though the visual-world paradigm has some limitations, it is still considered to be one of the best choice of doing research on word recognition and syntax processing in spoken language.

With the development of experimental equipment, if combined with ERP, the paradigm will show its power more effectively.

References

[1] Allopenna, Paul. D. , Magnuson, James. S. , & Tanenhaus, Michael. K. Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. Journal of Memory and Language, 1998,38:419-439.

[2] Berends, S. M., Brouwer, S M. & Sprenger, S. A. Eye tracking and the visual world paradigm https://www.researchgate.net/publication/281102556_Eye-Tracking_and _the_Visual_World_Paradigm, 2015.

[3] Chambers, C.G., Tanenhaus, M.K., Magnuson, J.S. Actions and Affordances in Syntactic Ambiguity Resolution.Journal of Experimental Psychology: Learning, Memory and Cognition. 2004, 30(3):687-696.

[4] Contemori, C., Carlson, M. and Marinis, T. On－line processing of English which－questions by children and adults: a visual world paradigm study. Journal of Child Language,2018, 45 (2): 415－441.

[5] Dahan, D., ＆Tanenhaus, M. K. Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. Journal of Experimental Psychology: Learning, Memory, and Cognition, 2004a, 30 (2):498-513.

[6] Dahan, D., ＆Tanenhaus, M. K. Looking at the rope when looking for the snake: Conceptually Mediated Eye Movements during Spoken-word Recognition. Psychonomic Bulletin and Review, 2004b, 12(3):453-459.

[7] Elman. J., & McClelland, J. Exploring lawful variability in the speech waveform. In S. Perkell & D.H. Klatt (Eds.), Invariance and variability in speech processing (pp. 360-385). Hillsdale,NJ: Erlbaum. 1986.

[8] Huettig, Falk., Rommers, Joost. & Meyer, Antje. S. Using the visual world paradigm to study language processing: A review and critical evaluation. Acta Psychologica, 2010, 137(2):151-171.

[9] Huettig, F., ＆ Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. Cognition, 2005, 96 (1):23-32.

[10] Liu, Meili. Eye Movement research on sound and form recognition of Mongolian language and Chinese of Mongolian College Students. Northwest Minzu University Master's Dissertation, 2018.

[11] Marslen-Wilson, W.D. Function and process in spoken word recognition －a tutorial review. In H. Bouma., & D.G. Bouwhuis (Eds). Attention & performance X. London: Lawrence Erlbaum. 1984.

[12] Marslen-Wilson, W.D., & Tyler, L. K. The temporal structure of spoken language understanding. Cognition, 1980,8:1-71.

[13] Marslen-Wilson, W.D., & Welsh, A. Processing interactions and lexical access during word-recognition in continuous speech. Cognitive Psychology, 1978, 10:29-63.

[14] Marslen-Wilson, W. D. Functional Parallelism in Spoken Word Recognition.Cognition, 1987, 25(1 -2):71-102.

[15] McClelland, J., & Elman, J. The TRACE model of speech perception. Cognitive Psychology, 1986, 18: 1-86.

[16] McQueen, J.M., & Viebahn, M.C. Tracking recognition of spoken words by tracking looks to printed words. Quarterly Journal of Experimental Psychology, 2007,60:661－671.

[17] Novick, J.M., Thompson-Schill, S.L., Trueswell, J.C. Putting lexical constraints in context into the Visual-World Paradigm. Cognition, 2008,107:850-903.

[18] Norris, D. Shortlist: A connectionist model of continuous speech recognition. Cognition, 1994, 52(3), 189-234.

[19] Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. The possible-word constraint in the segmentation of continuous speech. Cognitive Psychology, 1997,34(3), 191-243.

[20] Qiu, Lijing., Wang, Suiping., & Guan, Xin. Research on Visual-World Paradigm in Spoken Language Understanding, Journal of South China Normal University, 2009, 1(1): 130-136.

[21] Rommers, J., Meyer, A. S., Praamstra, P., ＆ Huettig, F. The contents of predictions in sentence comprehension: activation of the shape of objects before they are referred to. Neuropsychologia, 2013,51 (3):437-447.

[22] Shen, W., Qu, Q., ＆ Li, X. (2016). Semantic information mediates visual attention during spoken word recognition in Chinese: evidence from the printed-word version of the visual-world paradigm. Attention Perception and Psychophysics, 2016,78 (5):1-18.

[23] Snedeker, J. & Trueswell, J.C. The Developing Constraints on Parsing Decisions: The Role of Lexical-Biases and Referential Scenes in Child and Adult Sentence Processing. Cognitive Psychology, 2004, 49(3):238-299.

[24] Taft, Marcus. & Hambly,Gall. Exploring the Cohort Model of spoken word recognition. Cognition, 1986,22: 259-282.

[25] Tanenhaus, M. K., Spivey－Knowlton, M. J., Eberhard, K. M.,＆ Sedivy, J. C. Integration of visual and linguistic information in spoken language comprehension. Science, 1995, 268 (5217): 1632－1634.

[26] Yee, E., Sedivy, J.C. Eye movements to picture reveal transient semantic activation during spoken world recognition. Journal of Experimental Psychology: Learning, Memory and Cognition, 2006, 32(1):1-14.