

Exploring the frontiers of eye tracking research in language studies: a novel co-citation scientometric review

Vahid Aryadoust^a  and Bee Hoon Ang^b 

^aNational Institute of Education, Nanyang Technological University, Singapore; ^bSchool of Social Sciences- Psychology, Nanyang Technological University, Singapore

ABSTRACT

Eye tracking technology has become an increasingly popular methodology in language studies. Using data from 27 journals in language sciences indexed in the Social Science Citation Index and/or Scopus, we conducted an in-depth scientometric analysis of 341 research publications together with their 14,866 references between 1994 and 2018. We identified a number of countries, researchers, universities, and institutes with large numbers of publications in eye tracking research in language studies. We further discovered a mixed multitude of connected research trends that have shaped the nature and development of eye tracking research. Specifically, a document co-citation analysis revealed a number of major research clusters, their key topics, connections, and bursts (sudden citation surges). For example, the foci of clusters #0 through #5 were found to be perceptual learning, regressive eye movement(s), attributive adjective(s), stereotypical gender, discourse processing, and bilingual adult(s). The content of all the major clusters was closely examined and synthesized in the form of an in-depth review. Finally, we grounded the findings within a data-driven theory of scientific revolution and discussed how the observed patterns have contributed to the emergence of new trends. As the first scientometric investigation of eye tracking research in language studies, the present study offers several implications for future research that are discussed.

KEYWORDS

Burst; CiteSpace; cluster; co-citation; eye tracking; language; Scientometrics

Eye tracking is a process of detecting and measuring peoples' point of gaze (i.e., where they focus their visual attention) and eye movements (Holmqvist et al., 2011; Kredel, Vater, Klostermann, & Hossner, 2017; Rayner, 1998). In the field of linguistics, eye-trackers are generally used

CONTACT Vahid Aryadoust  vahid.aryadoust@nie.edu.sg  National Institute of Education, Nanyang Technological University, Singapore, Singapore.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

for measuring fixations (i.e., temporal pauses on a single location), saccades (i.e., rapid movement of the eyes from one fixation to another), and pupil size (Hepach, Vaish, & Tomasello, 2015; Suvorov, 2015; Vitu, 2005).

In recent years, there is a surge of interest in the application of eye tracking technologies in language studies. Some of these areas include the investigation of the effect of reading text captions or test items on listening performance (Aryadoust, 2019; Underwood, Jebbett, & Roberts, 2004; Yeldham, 2018), speaking skills (Yu et al., 2016), discourse or language processing (Altmann & Kamide, 2009), and syntactic processing (Frenck-Mestre, 2005). In addition, several reviews have synthesized eye tracking research in language studies and other fields including, psychology (Rayner, 1998), multimedia learning (Alemdag & Cagiltay, 2018), general education (Jarodzka, Holmqvist, & Gruber, 2017), reading times (Miller, 2015), music (Puurttinen, 2018), cognitive processes of drivers (Grüner & Ansorge, 2017). While such surveys of the literature are valuable in providing a coherent narration of the published research, they are subject to some limitations.

The first limitation of previous reviews lies in their breadth and scope (Chen et al., 2009). In these reviews, a narrow corpus of publications is typically summarized and segmented by the authors to investigate the state of research. Such focus occurs primarily due to limitations in cognitive capacity as it is time-consuming and laborious to read large corpora of papers, identify key ideas and links, and draw conclusions. According to Chen, Ibekwe-SanJuan, and Hou (2010, p. 1387), ‘essential tasks for interpretation such as categorization, summarization, synthesis, and integration are not only cognitively demanding but also inadequately supported’. Aside from the limited generalizability to the overall field of research (specialty) (Chen, 2004), such traditional review approaches often overlook the dynamics and connections between the articles, authors, and the journals in which the studies were published. The aforementioned information is critical towards understanding (i) the intellectual turning points within a specialty, (ii) the links between different specialties, and (iii) the progress of knowledge over time (Chen, 2004, 2013; Chen & Leydesdorff, 2014).

To address these gaps, the present study adopted Scientometrics to synthesize the available research on eye tracking in language studies from 1994 to 2018 (Chen & Song, 2017). Scientometrics or ‘the science of science’ (Fortunato et al., 2018) is a class of measurement methods that are used for investigating the development of underlying patterns in and the interconnections across various scientific fields (Nalimov & Mul’chenko, 1971). Additionally, Scientometrics also ‘relies on a broad

collection of quantitative methods, from descriptive statistics and data visualization to advanced econometric methods, network science approaches, machine-learning algorithms, mathematical analysis, and computer simulation, including agent-based modeling' (Fortunato et al., 2018, p. 1). The strength of scientometrics lies in its ability to categorize the disciplinary boundaries (Fortunato et al., 2018). This field is related to, but different from. Bibliometrics is the quantitative analysis of the metadata or the properties of the published literature (Fairthorne, 1969; Qiu, Zhao, Yang, & Dong, 2017; for further detail, see Glänzel & Schoepflin, 1994; Hood & Wilson, 2001). Bibliometrics has been used in various studies in applied linguistics (Ching, 2012; Lei & Liu, 2019), computer-assisted language learning (CALL; Jung, 2005), vocabulary acquisition (Meara, 2012), World Englishes (Arik & Arik, 2015), and second language (L2) writing (Arik & Arik, 2017). Up to the present point, there has not been any scientometric investigation of any topics in CALL and applied linguistics. According to Chen et al. (2010), the scientometric approach is increasingly employed in various scientific fields due to (1) the availability of large bibliographic corpora, such as Scopus and Web of Science (WoS; Bar-Ilan, 2008), (2) the availability of visualization and text-mining software packages, such as Many Eyes (IBM Corporation, NY, USA) and CiteSpace (Chen, 2016; Chen et al., 2010), and (3) the difficulty of analyzing vast quantities of data from various sources, especially the impossibility of finding co-citations using traditional review methods.

Unlike previous reviews that began with examining a selected number of publications in a specialty and drawing inferences from the results, in the present study, the relevant publications were firstly identified using Scopus and then segmented and synthesized using the CiteSpace software (Chen, 2016; Chen et al., 2010). The review of the influential trends within a specialty was presented only after the trends were identified.

Methods

Data source

In our experience, both Scopus and the Web of Science could return “false positives” or publications that were not closely relevant to search keywords. Thus, stringent inclusion criteria were used in the present study to create the dataset. First, the Scopus’s search engine was set to search for “eye tracking” in the titles, keywords, or abstracts of publications. The Scopus search was then streamlined by selecting the subject areas as Social Sciences, Arts and Humanities, Psychology, Computer Sciences, Health Professions, Neuroscience, Medicine, and Nursing.

Subsequently, another round of searches was conducted to narrow down the database to publications on reading, listening, speaking, writing, and language. Finally, the titles and abstracts of the suggested publications were read and only those that were closely related to language studies were retained. This process yielded a total of 363 publications ($n_{\text{paper}} = 358$; $n_{\text{book}} = 5$) that were published from 1994 to 2018. The data were subsequently downloaded from the Scopus depository into RIS formatted files.

Next, the “Import/Export” option in CiteSpace Version 5.3 (Chen, 2016; Chen et al., 2010) was used to convert the RIS formatted files to a CiteSpace-friendly format, with a successful conversion rate of $\sim 99\%$. Although the success rates vary in file conversions, a rate above 95% was considered reliable (Chen, 2013). From the 363 papers identified, CiteSpace identified 341 publications and 14,866 references ($n = 341 + 14,866 = 15,270$) to generate the networks. The remaining 22 publications were excluded from the analysis due to the absence of co-citations by the other publications in the network.

Data analysis

Descriptive statistics

A frequency analysis was performed to establish the number of eye tracking papers that were published annually, the names of the journals wherein the papers were published, and the names of the most productive authors, universities/institutes, and countries/regions wherein the authors were residing when the papers were published. Overall, there is an exponential trend in eye tracking research across language sciences since 1994 (Figure 1a). Overall, particularly after 2011 the number of publications on an annual basis were two-to-four times higher compared with the numbers in the 1990s and early 2000s.

Figure 1b highlights the most prolific authors, with S. Brown-Schmidt and M. J. Traxler at the top of the publication list (11 papers each), followed by M. K. Tanenhaus (8 papers), and M. J. Pickering, and P. Sturt (7 papers each). Additionally, there were seven authors with one paper each, 96 authors with two papers each, 28 authors with three papers each, 16 authors with four papers each, seven authors with five papers each, and one author with six papers. In terms of the universities and institutes that were involved in eye tracking language research, the Max Planck Institute for Psycholinguistics topped the list with 19 papers, while the University of Illinois at Urbana-Champaign, University of Glasgow, and University of Edinburgh were all tied in second position with 15 papers each (Figure 1c). Most papers (183 papers) were



Figure 1. Frequency analysis of the 363 publications on eye tracking in language studies.

published by authors from the USA, followed by the UK (71 papers), and Germany (50). China was the only Asian country among the top 10 countries. Other Asian countries were further down the list, including Japan and South Korea (8 papers each), Israel (3 papers), and Singapore and Taiwan (3 papers per each), along with South American countries like Chile (3 papers) and Mexico (1 paper).

Dual-map overlay

The CiteSpace software, Version 5.3, (Chen, 2016; Chen et al., 2010) was used to generate a dual-map overlay for investigating the current state of eye tracking research in language studies compared with other specialties. The dual-map overlay categorized the literature into two groups: (1) cited journals and (2) citing journals (i.e., the latter cited its references from the former). The strength of connections between these groups were visually represented and measured. The map generated in the present study superimposed the eye tracking papers and journal clusters that were often cited in language studies onto a base map derived from the WoS that comprised 10,000 journals (Chen & Leydesdorff, 2014).

Network visualization

To model the intellectual structure of the corpus, a network that included a series of published papers from 1994 to 2018 was firstly established. This process included investigating and visualizing the relevant journals, papers, and topics. A document co-citation analysis (DCA) and a journal co-citation analysis (JCA) that mapped a specialty domain ‘in terms of a time series of networks’ (Chen, 2016, p. 49) were performed

using a time-slicing technique. While the DCA detected impactful documents and publication clusters, the JCA was used to identify influential journals. According to Chen and Song (2017, p. 110):

The nature of a cluster is identified by the following aspects: a hierarchy of key terms in articles that cite the cluster (Tibély, Pollner, Vicsek, & Palla, 2013), the prominent members of the cluster as the intellectual milestones in its evolution and as the intellectual base of the specialty, recurring themes in the citing articles to the cluster to reflect the interrelationship between the intellectual base and the research fronts.

Quality control and impact

The qualities of the DCA and JCA networks, and the homogeneity of the detected clusters were measured using the modularity Q index and the average silhouette metric, respectively (Chen et al., 2010). The modularity Q index ranges between 0 and 1, with larger indices indicating higher reliability. The average silhouette metric ranges between -1 and 1, where values above 0 indicate better homogeneity.

In addition, influential publications and journals were determined by computing the burstness and betweenness for them (Kleinberg, 2002). A citation burst is a sudden surge of citations for a specific article and is indicated by a red ring around the node. According to Chen and Song (2017, p. 79), ‘a burst is an abrupt elevation of the frequencies [of citations] over a specific time interval’. Betweenness is a measure of influence that shows the degree to which publications or journals stand between each other. Publications with higher betweenness would have a higher influence on the network, because they connect more publications or journals and, accordingly, more information and paths pass through them (Chen, 2016).

Results

Generating the dual-map overlay

As illustrated in Figure 2, the multiple labels on the dual-map overlay were extracted from the journal titles representing the disciplines involved in the mapping. The citing journals are shown on the left side, and the cited journals on the right side. The arcs in the map represent the citation links starting from the citing journals on the left and pointing to the journals on the right, indicating how the citing journals are linked with the cited journals. The ovals represent the number of authors involved in publishing in the field and the number of relevant papers published. For example, the large oval on the left indicates that there are numerous authors and papers on eye tracking in language studies within the fields of ‘Psychology, Education, and Health’.

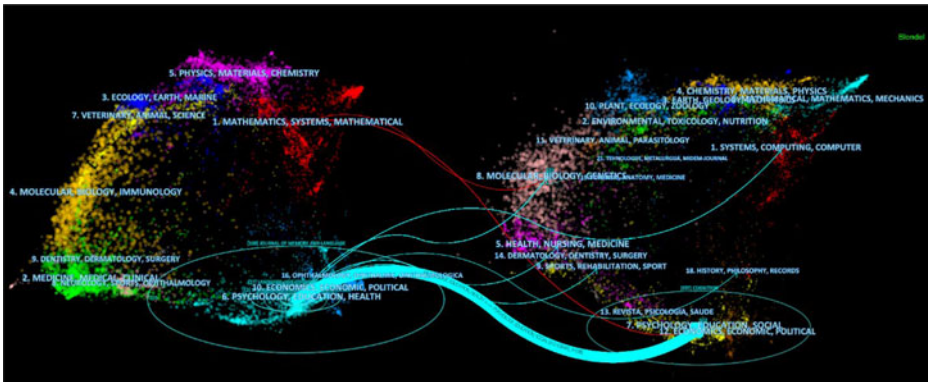


Figure 2. The dual-map overlay for eye tracking research.

The author-to-paper ratio is denoted by the width of the oval, while the height of the oval indicates the number of papers published.

The current view represents the z-score function and the significance of the connections is illustrated by the thickness of the connecting lines. For example, the thick line connecting ‘Psychology, Education, and Health’ on the left with ‘Psychology, Education, and Social [Sciences]’ on the right represents the most significant connections between eye tracking papers published in ‘Psychology, Education, and Health’ journals that have cited papers published in Psychology, Education, and Social journals. The line that connects ‘Psychology, Education, and Health’ to ‘Systems, Computer’ suggests that the former group of authors/papers cited the latter group of journals. A link was found between ‘Psychology, Education, and Health’ and ‘Molecular Biology, Genetics’, demonstrating a respectable connection between eye tracking research in language studies and these fields of science.

Document co-citation analysis

The modularity Q index and the average silhouette metric for the DCA network in the present study were 0.8141 and 0.4008, respectively, suggesting an acceptable level of reliability and homogeneity for the network. As shown in Figure 3, a total of 15 major clusters with labels were automatically extracted by CiteSpace Version 5.3 software based on the Latent Semantic Indexing (LSI) analysis (Deerwester, Dumais, Landauer, Furnas, & Harshman, 1990; smaller clusters were not displayed as they had little or no strength and connection).

Figure 4 presents the major clusters in the data on a horizontal line with the names of the most influential authors appearing below the line and the cluster labels appearing on the right side. The nodes with the

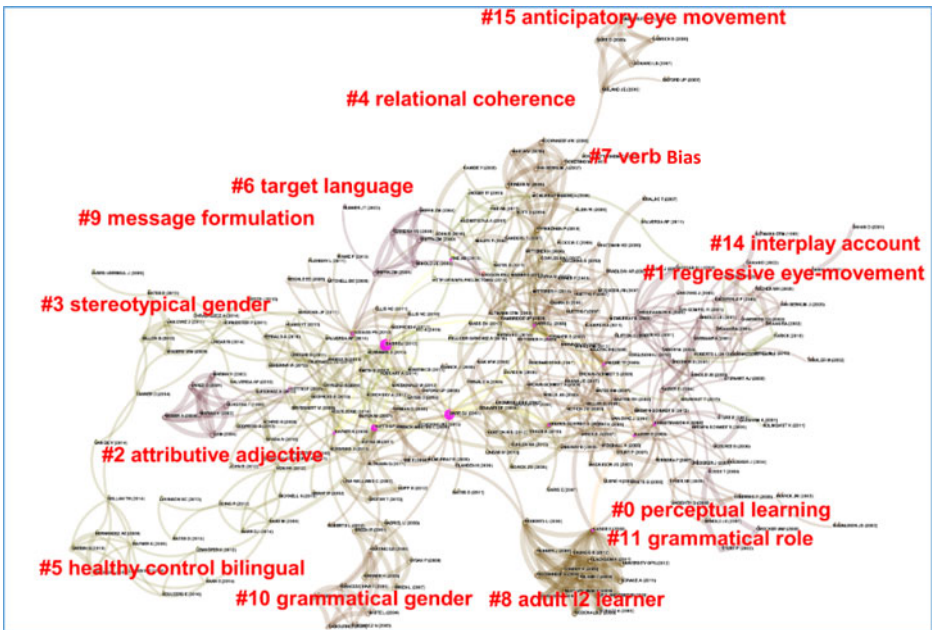


Figure 3. Hierarchical structure of the document co-citation analysis (DCA) network consisting of multiple clusters. (Modularity Q index = 0.8141; Average Silhouette = 0.4008).

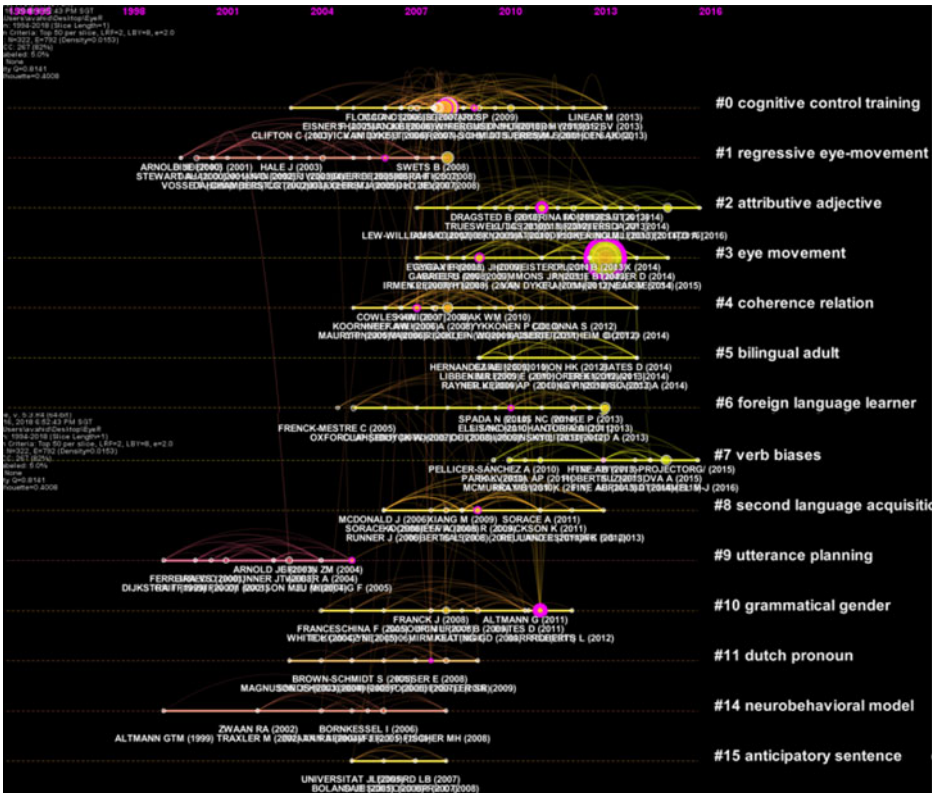


Figure 4. Visualization of clusters of the published research emerging from 1994 to 2018, using the Document Co-Citation Analysis (DCA). (Modularity Q index = 0.8141; Average Silhouette = 0.4008).

outer purple rings indicate high betweenness centrality, which is an index to measure strategic positions and the ability to connect different nodes in the network (Chen, 2016).

The length of each line represents the lifetime of the cluster (Chen, 2016). For example, the lifetime of cluster 0 was 2003–2013, whereas the lifetime of cluster 15 was much shorter (2004–2007). Cluster labels were generated by the text-mining and keyword analysis algorithm in CiteSpace software and included cognitive control training (0), regressive eye movement (2), attribute adjective (3) for example. There were multiple turning points in each cluster that serve as links between separate clusters. The largest turning point in the cluster ‘cognitive control training (0)’ is Baayen, Davidson, and Bates (2008) that was connected to clusters #4 and #11, indicating that this research paper influenced its own cluster as well as the eye tracking research on ‘coherence relations’ and ‘Dutch pronouns’ in other clusters. By contrast, some clusters such as ‘anticipatory sentence pronouns’ (15) and ‘utterance planning’ (9) had minimal connections with other clusters.

Document co-citation analysis: burst detection

A burst detection analysis (Kleinberg, 2002) was performed to identify the most influential or landmark publications in the field that had drawn researchers’ attention. Table 1 presents the 11 major bursts in citations, with the duration of each burst depicted on the right columns. There was a conspicuous drifting pattern in the data and this indicated that new research topics had emerged over time and the previously highly cited publications were gradually replaced by more current publications. For example, the citation burst of Rayner (1998) that has a strength or citation burst weight of 5.4931 lasted one year (2004–2005), whereas the citation burst of Kaiser and Trueswell (2004) lasted three years (2008–2011). These results suggested that these publications had been highly valuable for the specialty under investigation, especially during the beginning and end years of their bursts. Another notable finding from Table 1 is co-burst patterns, where two or more publications achieve the same magnitude of strength over the same time period. For example, the burst periods of Vosse and Kempen (2000) and Masson and Loftus (2003) overlap (2006–2008), although the former had higher strength. Notably, it took 1 (Knoeferle & Crocker, 2006) to 6 years (Rayner, 1998; Vosse & Kempen, 2000) for the publications to become a focal point and achieve citation bursts. Finally, Barr, Levy, Scheepers, and Tily (2013) and Bates, Machler, Bolker, and Walker (2015) are the most current citation bursts.

Table 1. Top 12 references with the strongest citation bursts.

References	Year	Strength	Begin	End
Rayner (1998)	1998	5.4931	2004	2005
Masson and Loftus (2003)	2003	5.4913	2006	2008
Griffin (2001)	2001	4.0679	2006	2007
Vosse and Kempen (2000)	2000	3.6469	2006	2008
Knoeferle and Crocker (2006)	2006	3.5867	2007	2010
Kaiser and Trueswell (2004)	2004	4.4489	2008	2011
Baayen et al. (2008)	2008	3.3308	2010	2012
Baayen et al. (2008)	2008	4.1324	2010	2012
Snedeker and Yuan (2008)	2008	3.6125	2011	2012
Barr et al. (2013)	2013	9.0187	2016	2018
Bates et al. (2015)	2015	6.4868	2017	2018

Document co-citation analysis: cluster characteristics

Table 2 presents the major clusters with the most citing published work per cluster (clusters #12 and #13 were excluded as they did not contain highly citing papers). The coverage column gives the number of references per cluster cited by the papers in their bibliography. For example, Novick, Hussey, Teubner-Rhodes, Harbison, and Bunting (2014) enjoyed a coverage index of 0.32 and this indicated that this paper cited 32% of the total references in this cluster. Alternatively, some publications were central to more than one cluster. The research from Andreu, Sanz-Torrent, and Trueswell (2013) that appeared in clusters #0 and #7 indicated a link across the distinct groups of publications (Chen, 2017). The content and scope of each cluster are further discussed in the ‘Discussion’ section.

Journal co-citation analysis

The modularity Q index and the average silhouette metric for the JCA network in the present study were 0.273 and 0.398 respectively and this suggested that the journal network had low degrees of homogeneity and connectedness. Table 3 presents a list of the top 24 journals in the field with strongest citation bursts. The earliest citation bursts in eye tracking research were found in *Cognitive Psychology* (1996–2003) and *Science* (1996–2009), while the more recent citation bursts occurred in *Behavior Research Methods* (2015–2018) and *Frontiers in Psychology* and *Perception* (2016–2018). Based on the strength statistics, the most influential journal in the eye tracking specialty in language sciences is *Frontiers in Psychology* (strength = 12.2012), followed by *Journal of Psycholinguistic Research* (strength = 6.6743), *Journal of Verbal Learning and Verbal Behavior* (strength = 6.3729), and *Journal of the Acoustic Society of America* (strength = 6.0572).

Discussion

In the present study, DCA and JCA, two co-citation analyses adopted from Scientometrics, were applied to mine a corpus of 363 publications

Table 2. Coverage and content of the clusters.

Coverage	Total reference	Representative authors ¹	Suggested title of the cluster
Cluster 0			
0.32	51	Novick et al. (2014)	Perceptual learning
0.28	15	Mitterer and Reinisch (2013)	
0.24	11	Brouwer et al. (2013)	
0.12	14	Trude et al. (2013)	
0.08	12	Yoon and Brown-Schmidt (2013)	
0.04	8	Andreu, Sanz-Torrent, and Trueswell (2013)	
0.04	12	Mak and Sanders (2013)	
Cluster 1			
0.5	44	Mitchell et al. (2008)	Regressive eye-movement
0.43	62	Kaiser and Trueswell (2004)	
0.07	26	Crocker et al. (2010)	
Cluster 2			
0.36	31		Attributive adjective
0.36	2	Tribushinina and Mak (2016)	
0.07	13	Andringa and Curcic (2015)	
0.07	9	Hopp (2015)	
0.07	2	Ito et al. (2018)	
Cluster 3			
0.3	11	Esaulova and von Stockhausen (2015)	Stereotypical gender
0.3	4	Reali, Esaulova, and von Stockhausen (2015)	
0.2	1	Miwa and Dijkstra (2017)	
0.2			
0.1			
Cluster 4			
0.69	12	Mak and Sanders (2013)	Relational coherence & attention
0.08	40	Winke et al. (2013)	
Cluster 5			
1.2	2	Ratiu and Azuma (2017)	Healthy control bilingual
0.6	1		
Cluster 6			
0.67	40	Winke et al. (2013)	Target language in speech perception
0.17	19	Roberts and Siyanova-Chanturia (2013)	
Cluster 7			
0.44	1	Ryskin et al. (2018)	Verb bias
0.22	8	Andreu, Sanz-Torrent, and Trueswell (2013)	
0.11	19	Farris-Trimble & McMurray (2013)	
Cluster 8			
0.5	6	Kim et al. (2015)	Adult L2 learner
Cluster 9			
0.38	84	Brown-Schmidt and Tanenhaus (2006)	Message formulation
0.25	85	Blumenfeld and Marian (2007)	
0.25	48	Kaushanskaya and Marian (2007)	
Cluster 10			
0.55	72	Grüter et al. (2012)	Grammatical gender
Cluster 11			
0.5	8	Kaiser (2011)	Reference resolution
Coverage 14			
0.57	26	Crocker et al. (2010)	Interplay account
Coverage 15			
0.75	8	Andreu, Sanz-Torrent, Olmos, et al. (2013)	Anticipatory eye movement

¹The publications reported in this table constitute the most prominent bursts per cluster.

on eye tracking in language studies from Scopus. Overall, 15 clusters of research and 24 influential journals were identified. The main characteristics of each cluster from the DCA and their contributions to eye tracking research in language studies are firstly discussed below, followed by results from the JCA and the implications of the findings. The synthesis of clusters #5 through #15 is provided in [Appendix](#).

Table 3. Top 24 cited journals with the strongest citation bursts.

Cited Journals	Strength	Begin	End
Cognitive Psychology	5.0739	1996	2003
Science	3.5301	1996	2009
Journal of Verbal Learning and Verbal Behavior	6.3729	1998	2004
Journal of Psycholinguistic Research	6.6743	2000	2009
Linguistics and Philosophy	3.5506	2002	2011
Canadian Journal of Experimental Psychology	5.0581	2002	2008
Journal of Semantics	3.3012	2002	2010
Journal of Experimental Psychology	3.6641	2004	2011
Computational Linguistics	3.5282	2007	2011
Behavioral and Brain Sciences	3.3897	2007	2010
Memory and Cognition	4.8538	2008	2010
Natural Language Semantics	3.5776	2010	2012
Journal of the Acoustical Society of America	6.0572	2011	2014
Journal of Speech and Hearing Research	4.6141	2011	2013
Journal of Linguistics	3.9881	2011	2014
Journal of Phonetics	5.0446	2011	2013
Perception & Psychophysics	4.6377	2012	2013
Applied Linguistics	5.496	2014	2015
Modern Language Journal	4.4467	2014	2015
Current Directions in Psychological Science	4.322	2014	2016
Behavior Research Methods	4.7429	2015	2018
Frontiers in Psychology	12.2012	2016	2018
Perception	3.3389	2016	2018

Document co-citation analysis (DCA)

Cluster 0: Perceptual learning (through listening)

Publications in this cluster applied eye tracking to investigate (1) perceptual learning by measuring the role of cognitive control training and executive function when recovering from miscomprehension in reading (Novick et al., 2014) and (2) recovery from ambiguity in consonants (Mitterer & Reinisch, 2013) and sentences (Brouwer, Mitterer, & Huettig, 2013). Novick et al. (2014) found that executive function (EF) training significantly improved trainees' eye movement patterns when revising miscomprehension. They argued that 'when confronted with new evidence that conflicts with developing interpretations, readers who undergo EF training [...] spend less time regressing to earlier material in order to recover successfully from their misanalysis, effectively gathering information more quickly to arrive at the correct sentence meaning' (Novick et al., 2014, p. 203). Novick et al. (2014) emphasized the utility of eye movement measurements and suggested that regressive (backward) saccades during reading initiates the syntactic ambiguity rectification process. In another study, Mitterer and Reinisch (2013) who investigated listeners' adjustments to speakers' idiosyncratic pronunciation of fricatives found that listeners detected ambiguous fricatives simultaneously as they were processing them. Mitterer and Reinisch (2013) reported that, after hearing the target fricative, it took listeners 200 ms to plan (McMurray, Clayards, Tanenhaus, & Aslin, 2008), although other

researchers have proposed 100 ms as the window between speech signals and planning (Altmann, 2011).

In another study, Brouwer et al. (2013, p. 520) reported that ‘participants used supportive discourse contexts to predict the upcoming target word in the canonical-target condition but not in the reduced-target condition’. By contrast, after ambiguous (reduced) linguistic forms were heard, the supporting effect of context reversed in favor of the reduced-target forms. These findings suggest the presence of anticipation, which, were consistent with the findings of Mitterer and Reinisch (2013) in this cluster. Brouwer et al. (2013) argued that when exposure to sample speech is adequately long, speaker adaptation and sufficient discourse context could support the processing of canonical and reduced forms.

Similar to Brouwer et al. (2013), Trude, Tremblay, and Brown-Schmidt (2013) investigated whether listeners can overcome sources of variability heard during speeches if they are only exposed to a sample of the speakers’ speech. In multiple experiments, they found that: (1) listeners fixated on the wrong words (competitors) more often than on the target words that had been phonetically assimilated to the competitors, but that they accommodated speakers’ accents over time, which is consistent with Brouwer et al. (2013); (2) simplification of phonotactic regularity of a set of accented words did not help listeners process foreign-accented words; (3) listeners performed unexpectedly better when they heard speech including non-shifted sounds as opposed to when they listened to the discourse including shifted sounds; (4) when a vowel was shifted in a word without competitors (e.g., a long vowel replaced by a short one, creating a non-sense word), listeners learned the anomaly faster (perceptual learning); however, learning did not advantage them significantly compared to the situation with competitors; and (5) foreign-accented shifts to vowels in a string of discourse are more difficult to process than the same shift made by a native speaker. In an earlier study, Trude and Brown-Schmidt (2012), who investigated male and female English speakers with Oregon and Chicago accents, had found that participants learned information about other forms of speech quickly and integrated this new knowledge to ensure smooth communication.

Finally, Yoon and Brown-Schmidt (2013) found that speakers considered the history of conversation using referential expression, although there was no evidence showing that listeners used lexical differentiation. Other main findings in this cluster are the lack of significant differences between children with specific language impairment (SLI) and typically developing children or adults in retrieving verb meaning and anticipating upcoming objects. However, the semantic features of object nouns

modulate ‘anticipatory eye movements’, as discussed in cluster #15, too (Andreu, Sanz-Torrent, & Trueswell, 2013). Other relevant publications within this cluster include Holsinger (2013), Julie, Millotte, Posada, and Rizzi (2013), Kush, Lidz, and Phillips (2015), Roembke and McMurray (2016), Tremblay and Spinelli (2013), and von der Malsburg and Vasishth (2013).

Cluster 1: Regressive eye movement

In this cluster, Mitchell, Shen, Green, and Hodgson (2008), Crocker, Knoeferle, and Mayberry (2010), Kaiser and Trueswell (2004), Knoeferle and Crocker (2006), and Tribushinina and Mak (2016) were identified as the most influential publications. Mitchell et al. (2008, p. 266) found that, in linguistically manipulated sentences, the physical formatting of text significantly influences the ‘landing site of the first regression’ from keywords, and that in sentences wherein the position of ambiguous words was linguistically fixed, readers’ regressive movements tended to be influenced by linguistic cues in the text. Regressive eye movements are also a main measurement indicator in other studies. Crocker et al. (2010) argued that it is natural for humans to reconcile the languages they experience with other aspects of the world around them, such as their world knowledge or other sensory experiences. Relatedly, Knoeferle and Crocker (2006) developed the Coordinated Interplay Account (CIA) to explain the phenomenon where incremental interpretation of speech drives visual attention towards anticipated or mentioned objects or events. Such an incremental view of language comprehension is supported by various studies. For instance, Tribushinina and Mak (2016) found that children gazed at the correct noun picture upon hearing the adjective as much as adults did, indicating similar tendency for predicting the noun upon hearing the prenominal adjective. Corley (2010) simultaneously presented visual scenes and a stimulus sentence to 24 university students and found that, upon hearing the verb (e.g., ring), and participants were able to fixate on the correct agent in the scene (e.g., a bell). More importantly, when the sentence was repaired (e.g., ‘the baby will ring—uh, kick the bell’), eye fixations on the agent changed prior to the noun, indicating constant update of language comprehension as the sentence unfolded.

While these studies directly measured eye movements, including regressions, as speech progressed, other studies added to the compelling evidence for CIA by examining how speech misinterpretation is accompanied by a coping mechanism in the form of visual attention. For instance, researchers reported gaze transfers and regressive eye movements towards the speakers’ mouth when listening to stuttered speeches,

indicating difficulty in understanding and interpreting speeches and, hence, a requirement for a coping mechanism to better comprehend the input (Bowers, Crawcour, Saltuklaroglu, & Kalinowski, 2010; Zhang & Kalinowski, 2012). Regressive eye-movement is also a strategy observed in coping with impaired comprehension, whereby one re-reads earlier materials (Mitchell et al., 2008).

With a better understanding of the relationship between language processing and visual attention via eye tracking, other studies in this cluster have examined the factors mediating these two processes (Nazir, Ben-Boutayab, Decoppet, Deutsch, & Frost, 2004; Crocker et al., 2010; Tribushinina & Mak, 2016). The immediate environment of a child as he/she learns to ground concepts with visual referents during language acquisition is of paramount importance. Language expectations, such as discourse and stress expectations, can affect a child's visual attention (Breen & Clifton, 2011; Roland, Maurer, O'Meara, & Yun, 2012). Various disorders can affect this synchrony, such as apraxia (Myung et al., 2010) or SLI (Andreu, Sanz-Torrent, Olmos, & MacWhinney, 2013). Other relevant publications in this cluster include Caplan (2010), Foucart and Frenck-Mestre (2012), Humphreys, Mirković, and Gennari (2016), Myung et al. (2010), and Stites, Federmeier, and Christianson (2016).

Cluster 2: Attributive adjectives

Studies in this cluster adopted certain grammatical structures, notably attributive adjectives (i.e., those that are presented before a noun), to investigate the phenomenon in which incremental interpretation of speech drives visual attention towards anticipated or mentioned objects or events, which in turn affects subsequent comprehension (Knoeferle & Crocker, 2006). In other words, language is largely affected by external visual stimuli. For instance, Rader and Zukow-Goldring (2012) found that infants learn better when a word is paired with a caregiver's dynamic gestures (e.g., pointing).

Another related phenomenon is one's tendency to make predictions as verbal language is heard, possibly for more efficient comprehension processing (Andringa & Curcic, 2015). Several studies supported this, including Tribushinina and Mak (2016) who employed the visual world paradigm with 21 Dutch speaking children (mean age = 3.2 years) in which they were shown a picture combination and later given a noun phrase including an attributive adjective (e.g., the heavy stone). Tribushinina and Mak found that children gazed at the correct noun picture upon hearing the attributive adjective as much as adults, indicating

comparable proficiency for predicting the noun upon hearing the pre-nominal attributive adjective.

Based on these findings, we found an increasing focus on how and whether the results hold for multilingual individuals. Several studies have examined bilingual individuals who knew at least two languages, a primary or level one language (L1) and a secondary or level two language (L2). Ito, Pickering, and Corley (2018) recently found that such individuals are equally capable in predicting the upcoming words across L1 and L2 languages. It was found that both L1 and L2 speakers gazed at the target object more than the distractors before they heard the target word. However, L1 speakers responded quicker, suggesting greater predictability sensitivity, possibly due to the limited resources associated with L2 processing. Hopp (2015) found an array of individual differences in L2 language processing due to differing efficiency integrating grammatical information, as well as differing capabilities in sub-categorization, lexical processing, and recruitment of morphosyntactic knowledge. Other relevant publications in this cluster include Beyer and Hudson-Kam (2009), Grassman and Tomasello (2010), Ito, Jincho, Minai, Yamane, and Mazuka (2012), and Wilkinson and Light (2011).

Cluster 3: Stereotypical gender

The label selected for this cluster is based on Esaulova and von Stockhausen (2015) and Reali, Esaulova, and von Stockhausen (2015) who investigated gender as a prominent feature across languages. Using gender as a main variable, studies in this cluster are primarily grounded by a multimodality framework that views communication as taking various forms including gestures, words, gazes, postures, etc. (Hiippala, 2012). Accordingly, communication requires a combination of these stimuli for holistic understanding. There is an array of studies supporting the multimodality of language and communication. For instance, Kuperman, Bertram, and Baayen (2010) examined the effects of word length on comprehension and processing. Suffix length was found to modulate word frequency and reading time. Another physical property of words—vertical or horizontal reading—was found to affect lexical processing. Specifically, Miwa and Dijkstra (2017) found that vertical reading resulted in a longer fixation duration and, hence, longer reading time—possibly due to lower visual acuity in the vertical direction. Lastly, physical gestures play a large part in communication, with Rader and Zukow-Goldring (2012) finding that infants learn better when a word is paired with a caregiver's dynamic gesture (e.g., pointing). (Rader and Zukow-Goldring's paper appeared in clusters 2, 3, and 4).

Besides the physical properties of language, various technical aspects affect comprehension and communication. These properties include prosodic stress (Zhou, Su, Crain, Gao, & Zhan, 2012), the use of scalar adjectives (Wolter, Gorman, & Tanenhaus, 2011), intentional verbs (e.g., ‘wanted a beer’), pronominalization, and subjecthood (Kaiser, 2011). Most prominently, Esaulova and von Stockhausen (2015) argued for an effect of a gender-based prominence feature on language processing that helps individuals organize information hierarchically, where higher ranked arguments are likely to be interpreted first as agents. Esaulova and von Stockhausen argued for stereotypical gender or role nouns to be considered as a prominence feature likely to affect subsequent comprehension. Taken together, these studies suggest that language comprehension and processing interact with various technical properties of language, thereby emphasizing the multimodality of language. This cluster also includes Brandt-Kobele and Hohle (2010), Hiippala (2012), Lowder and Gordon (2012), and Schmiedtova (2011),

Cluster 4: Relational coherence and attention

Two main topics covered in this cluster were: coherence and attention. When researching L2 language learning, eye movement recording is one of the richest records of cognitive processing of textual information (Winke, Sydorenko, & Gass, 2013). Winke et al. (2013) examined how the use of captions can improve learners’ generation of coherence and comprehension of a secondary language. In this study, Mandarin language learners were found to attend to captions more than videos, possibly because Mandarin, being non-alphabetical, is more difficult to process and thus requires one to focus attention on a single modality to create relational coherence in the stimuli. The role of coherence was further investigated by Mak and Sanders (2013), who showed that causal coherence is established even at the nascence of a connector. This finding indicates cause-effect relationships and that readers’ ‘causal expectations’ can differentially affect processing and comprehension of relational and referential information in texts. Mak and Sanders further argued that causal expectations may aid the processing of relational coherence but only if the readers were able to establish cause-effect relations between sentences.

Coherence building is also indispensable to the studies that investigated comprehension in cluster 4. Lansing and McConkie (1999) found that listeners focus on differing regions of the face when making judgements about different aspects of speech. Specifically, listeners focused on the upper face region when deciphering intonation and sentence stress, possibly because intonation and changing pitch to produce sentence

stress usually involve eyebrow motions. This finding highlights the need for visual attention to direct speech comprehension that demands the generation of coherence in speech. In another study, Godfroid, Boers, and Housen (2013) reported a direct and positive relationship between the amount of attention and learning. Learners were able to retain more words on an unannounced post-test when they fixated on these words longer while reading.

Given its effect on language acquisition, attention has become an important topic in secondary language learning research. Jackson, Dussias, and Hristovia (2012) examined if learners attend to L1 and L2 inputs differently. They recruited 25 intermediate L2 learners of German and instructed them to read a number of German sentences. These learners were found to spend more time processing subject-first than object-first sentences, regardless of the position of the verb. Jackson et al. (2012) argued that this may be due to a general preference for object-first sentences in L2 learners as opposed to German native speakers. Furthermore, the results indicate a tendency to attend to the agent—a semantic property—more than to the grammatical form, which is the location of verb. Taken together, these studies reveal the varied ways that attention can be manipulated to achieve better acquisition of L2 languages.

Journal co-citation analysis

The JCA identified *Science* as the journal with the longest burst duration (13 years) followed by *Journal of Psycholinguistic Research* and *Linguistics and Philosophy* (9 years), and *Journal of Semantics* (8 years). The burst period of these journals ended before 2012, while journals with the most recent high burstness are *Behavior Research Methods*, *Frontiers in Psychology*, and *Perception*, which, among other subjects, focus on language research—for example, *Frontiers in Psychology* includes a specific section devoted to language sciences and research and a special topic on language assessment. Interestingly, the burst of some of the influential journals in (applied) linguistics such as *The Modern Language Journal* and *Applied Linguistics* ended before 2016, indicating that the impact of eye tracking research published in these journals had a downward trend over the years.

Implications of the study

The study advances previous bibliometric research in CALL and applied linguistics (e.g., Arik & Arik, 2015, 2017; Ching, 2012; Jung, 2005; Lei &

Liu, 2019; Meara, 2012) in three important ways. First, adopting text-mining and clustering tools from Scientometrics allowed us to identify interdependencies of the eye-tracking specialty in language studies with other fields. The dual-map overlay showed that the field of eye tracking research in language sciences has forged the strongest connections with psychology and social sciences research, as well as some connections with systems, computers, medicine, genetics, and molecular biology. While strong connections indicate a well-established and perhaps routinized relationship among fields, the weak links could herald the emergence of interdisciplinary research or the shift of researchers' attention from within the field to beyond. This pattern could indicate the start of a gradual paradigm shift (Kuhn, 1962) and likely a new conceptualization stage (Chen & Song, 2017). Whereas the mere citation of papers from other disciplines cannot be regarded as a corroboration of paradigm shifts, it could be viewed as evidence that a group of language researchers have taken aspiration from other fields, which is a necessary (but insufficient) ingredient of a paradigm shift.

Second, there have been several distinct yet interconnected research clusters in eye tracking research since 1994, with multiple bursts that are regarded as tipping points of each cluster. Identifying latent chronological patterns in this field, for the first time, allows for exploring how the field has evolved over the years. The evolutionary stages of the specialty are documented by the progression of the clusters over time. The fact that several publications in each cluster have adopted or drawn upon theoretical frameworks from other clusters suggests that the specialty of eye tracking in language studies has reached what Chen and Song (2017) called the 'instrument building stage', which is characterized by innovation of new methods and tools, or perhaps the stage of instrument application, which is characterized by the application of new tools and/or methods in research to generate new insights and discover anomalies. In addition, the end date of the clusters can be regarded as knowledge codification (characterized by 'low productivity of knowledge' and routinization of research; Chen & Song, 2017, p. 52) where cluster-specific productions of knowledge and discovery diminishes and eventually ends (Chen & Song, 2017). Accordingly, cluster #2 (Attributive Adjectives) and clusters #7 (Verb Bias) can be viewed as the current frontiers of eye tracking research in language studies, as they both continue to thrive. Cluster #7 seems to be in the third stage of a scientific revolution, which is the application of the instruments and frameworks in other fields. According to Chen (2017, p. 12), '[i]f the methodologies appear in a cluster of knowledge domains external to information science [...], then [they would] consider [that] the specialty has reached Stage III—tools developed by the specialty are applied to other subject

domains'. Nevertheless, the pattern of transfer of theoretical frameworks from cluster #7 to other clusters is intra-disciplinary and understanding the potential impact of this cluster on other relevant fields warrants further investigation.

Third, several publications that traversed across different clusters were identified. These 'extr[a]ordinary or unexpected connections across distinct clusters' were important, since they indicate the significant contributions of one or more publications across the specialty (Chen, 2017, p. 30). This could be due to the 'intellectual base' of these publications (the cited publications of the cluster) being the 'research front' (the topics addressed in the clusters) of other clusters (Chen, 2017) or vice versa. Such publications have considerable transformative potential and their methodology and frameworks should be examined closely. Together, the three implications of the study establish a framework for future research in scientometric and bibliometric research in CALL and applied linguistics.

Conclusion

In summary, the scientometric approach used in the present study has provided evidence for the dynamicity of eye tracking research in language studies. Adopting visualization and co-citation techniques enabled us to map the frontiers of the specialty, interconnections between publications in the corpus used, and the prominent research trends. An interesting research agenda for the future would be to investigate the evolution of eye tracking across all relevant specialties. This would constitute a large-scale study that would shed light on the progress and decline of impactful research trends across specialties.

Acknowledgements

The authors wish to thank Ng Li Ying, Stacy Foo, and Brigit Sebastian for their suggestions and comments on earlier drafts of this article.

Funding

This study was supported by the National Institute of Education of Nanyang Technological University, Singapore (grant number = RI 2/16 VSA) and Paragon Testing Enterprises (Paragon Research Grants).

Notes on contributors

Vahid Aryadoust is an Assistant Professor of language assessment literacy at the National Institute of Education of Nanyang Technological University, Singapore. He has

led a number of language assessment research projects funded by, for example, the Ministry of Education (Singapore), Michigan Language Assessment (USA), Pearson Education (UK), and Paragon Testing Enterprises (Canada), and published his research in, for example, *Language Testing*, *Language Assessment Quarterly*, *Assessing Writing*, *Educational Assessment*, *Educational Psychology*, and *Computer Assisted Language Learning*. He has also (co)authored a number of book chapters and books published by Routledge, Cambridge University Press, Springer, Cambridge Scholar Publishing, Wiley Blackwell, etc. He is a member of the Advisory Board of multiple international journals and has been awarded the Intercontinental Academia Fellowship (2018–2019), which is an advanced research program launched by the University-Based Institutes for Advanced Studies. His most recent publication is a two-volume book published by Routledge (<https://www.routledge.com/products/search?author=Vahid%20Aryadoust>).

Ang Bee Hoon is a student in psychology at the School of Social Sciences of Nanyang Technological University of Singapore. She has acted as a research assistant in multiple research studies that span across the field of neuroscience, cognitive psychology, and language. She has also conducted various research studies in her seven years of Psychology study in her education institutes and has collaborated with governmental agencies of Singapore such as the Dyslexic Association of Singapore. Some of Bee Hoon's areas of interest include eye tracking, brain-imaging, and content analysis.

ORCID

Vahid Aryadoust  <http://orcid.org/0000-0001-6960-2489>

Bee Hoon Ang  <http://orcid.org/0000-0002-1997-5292>

References

- Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, 125, 413–428. doi:10.1016/j.compedu.2018.06.023
- Altmann, G. T. M. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychologica*, 137(2), 190–200. doi:10.1016/j.actpsy.2010.09.009
- Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye-movements and mental representation. *Cognition*, 111(1), 55–71. doi:10.1016/j.cognition.2008.12.005
- Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2013). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. *Applied Psycholinguistics*, 34(1), 5–44. doi:10.1017/S0142716411000592
- Andreu, L., Sanz-Torrent, M., Olmos, J. G., & MacWhinney, B. (2013). The formulation of argument structure in SLI: An eye-movement study. *Clinical Linguistics & Phonetics*, 27(2), 111–133. doi:10.3109/02699206.2012.751623
- Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing: Evidence from differential object marking acquisition. *Studies in Second Language Acquisition*, 37(2), 237. doi:10.1017/S0272263115000017

- Arik, B. T., & Arik, E. (2015). World Englishes from a citation index perspective. *EILJ*, 10(1), 1–19. Retrieved from <https://www.elejournals.com/1200/eilj-2015/eilj-volume-10-issue-1-may-2015/>
- Arik, B. T., & Arik, E. (2017). Second Language Writing” publications in Web of Science: A bibliometric analysis. *Publications*, 5, 4. doi:10.3390/publications5010004
- Aryadoust, V. (2019). Dynamics of item reading and answer changing in two hearings in a computerized while-listening performance test: An eye-tracking study. *Computer Assisted Language Learning (Learning)*, 1. doi:10.1080/09588221.2019.1574267
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. doi:10.1016/j.jml.2007.12.005
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.
- Bar-Ilan, J. (2008). Which h-index? A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257–271. doi:10.1007/s11192-008-0216-y
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. doi:10.1016/j.jml.2012.11.001
- Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01
- Bergmann, C., Paulus, M., & Fikkert, P. (2012). Preschoolers’ comprehension of pronouns and reflexives: the impact of the task. *Journal of Child Language*, 39(4), 777–803. doi:10.1017/S0305000911000298
- Beyer, T., & Hudson-Kam, C. L. (2009). Some cues are stronger than others: The (non)-interpretation of third person present -s as a tense marker by 6- and 7-year-olds. *First Language*, 29(2), 208–227. doi:10.1177/0142723708101678
- Blumenfeld, H., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, 22(5), 633–660. doi:10.1080/01690960601000746
- Bowers, A. L., Crawcour, S. C., Saltuklaroglu, T., & Kalinowski, J. (2010). Gaze aversion to stuttered speech: A pilot study investigating differential visual attention to stuttered and fluent speech. *International Journal of Language and Communication Disorders*, 45, 133–144.
- Brandt-Kobe, O. C., & Hohle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua*, 120(8), 1910–1925. doi:10.1016/j.lingua.2010.02.008
- Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, 64(2), 153–170. doi:10.1016/j.jml.2010.11.001
- Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics*, 34(3), 519–539. doi:10.1017/S0142716411000853
- Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54, 592–609. doi:10.1016/j.jml.2005.12.008
- Canestrelli, A., Mak, W., & Sanders, T. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye-movements. *Language and Cognitive Processes*, 28, 1394–1413. doi:10.1080/01690965.2012.685885

- Caplan, D. (2010). Task effects on BOLD signal correlates of implicit syntactic processing. *Language and Cognitive Processes*, 25(6), 866–901. doi:10.1080/01690961003672447
- Chen, C., & Leydesdorff, L. (2014). Patterns of connections and movements in dual-map overlays: A new method of publication portfolio analysis. *Journal of the Association for Information Science and Technology*, 65(2), 334–351. doi:10.1002/asi.22968
- Chen, C. (2004). Searching for intellectual turning points: Progressive Knowledge Domain Visualization. *Proceedings of the National Academy of Sciences*, 101(Supplement 1), 5303–5310. doi:10.1073/pnas.0307513100
- Chen, C. (2013). *Mapping scientific frontiers: The quest for knowledge visualization* (2nd ed.). London: Springer.
- Chen, C. (2016). *CiteSpace: A practical guide for mapping scientific literature*. Hauppauge, NY: Nova Publishers.
- Chen, C. (2017). Science mapping: A systematic review of the literature. *Journal of Data and Information Science*, 2(2), 1–40. doi:10.1515/jdis-2017-0006
- Chen, C., & Song, M. (2017). *Representing scientific knowledge: The role of uncertainty*. New York: Springer.
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3(3), 191–209. doi:10.1016/j.joi.2009.03.004
- Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis. *Journal of the American Society for Information Science and Technology*, 61(7), 1386–1409. doi:10.1002/asi.21309
- Ching, G. (2012). 40 years of language studies: A Bibliometric study of research articles in Taiwan. *Asian EFL Journal*, 64, 53–74.
- Corley, M. (2010). Making predictions from speech with repairs: Evidence from eye movements. *Language and Cognitive Processes*, 25(5), 706–727. doi:10.1080/01690960903512489
- Coventry, K. R., Lynott, D., Cangelosi, A., Monrouxe, L., Joyce, D., & Richardson, D. C. (2010). Spatial language, visual attention, and perceptual simulation. *Brain and Language*, 112(3), 202–213. doi:10.1016/j.bandl.2009.06.001
- Crocker, M. W., Knoeferle, P., & Mayberry, M. R. (2010). Situated sentence processing: The coordinated interplay account and a neurobehavioral model. *Brain and Language*, 112, 189–201. doi:10.1016/j.bandl.2009.03.004
- Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391–407. doi:10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9
- Ellert, M. (2013). Resolving ambiguous pronouns in a second language: A visual-world eye-tracking study with Dutch learners of German. *International Review of Applied Linguistics in Language Teaching*, 51(2), 171–197. doi:10.1515/iral-2013-0008
- Esaulova, Y., & von Stockhausen, L. (2015). Cross-linguistic evidence for gender as a prominence feature. *Hypothesis and Theory*, 6, 1356.
- Evitts, P. M., & Gallop, R. (2011). Objective eye-gaze behavior during face-to-face communication with proficient alaryngeal speakers: A preliminary study. *International Journal of Language & Communication Disorders*, 46(5), 535–549.
- Fairthorne, R. A. (1969). Empirical hyperbolic distributions (bradford-zipf-mandelbrot) for bibliometric description and prediction. *Journal of Documentation*, 25(4), 319–343. doi:10.1108/eb026481

- Farris-Trimble, A., & McMurray, B. (2013). Test-retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. *Journal of Speech, Language, and Hearing Research*, 56(4), 1328–1345.
- Farmer, T. A., Cargill, S., Hindy, N., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer-mouse trajectories suggest parallel syntactic processing. *Cognitive Science: A Multidisciplinary Journal*, 31(5), 889–909. doi: [10.1080/03640210701530797](https://doi.org/10.1080/03640210701530797)
- Fernald, A., Zangl, R., Thorpe, K., Hurtado, N., & Williams, C. (2006). *Learning to listen ahead in English and Spanish: Infants use multiple linguistic and non-linguistic cues in online sentence interpretation*. Paper presented at the Online Methods in Children's Language Processing Workshop, New York, NY.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... Barabási, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. doi:[10.1126/science.aao0185](https://doi.org/10.1126/science.aao0185)
- Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66(1), 226–248. doi:[10.1016/j.jml.2011.07.007](https://doi.org/10.1016/j.jml.2011.07.007)
- Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research*, 21(2), 175–198. doi:[10.1191/0267658305sr257oa](https://doi.org/10.1191/0267658305sr257oa)
- Glänzel, W., Schoepflin, U. (1994). Little Scientometrics, big Scientometrics ... and beyond? *Scientometrics*, 30(2-3), 375–384. doi:[10.1007/BF02018107](https://doi.org/10.1007/BF02018107)
- Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, 35(3), 483–517. doi:[10.1017/S0272263113000119](https://doi.org/10.1017/S0272263113000119)
- Grassman, S., & Tomasello, M. (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics*, 42, 3098–3105. doi:[10.1016/j.pragma.2010.04.019](https://doi.org/10.1016/j.pragma.2010.04.019)
- Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, 99(3), 272–288.
- Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82(1), B1–B14. doi:[10.1016/S0010-0277\(01\)00138-X](https://doi.org/10.1016/S0010-0277(01)00138-X)
- Grüner, M., & Ansorge, U. (2017). Mobile eye tracking during real-world night driving: A selective review of findings and recommendations for future research. *Journal of Eye Movement Research*, 10(2), 1–18. doi:[10.16910/jemr.10.2.1](https://doi.org/10.16910/jemr.10.2.1)
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem?. *Second Language Research*, 28(2), 191–215. doi:[10.1177/0267658312437990](https://doi.org/10.1177/0267658312437990)
- Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: Posture, gaze, and pupil dilation. *Frontiers in Psychology*, 6, 858. doi:[10.3389/fpsyg.2015.00858](https://doi.org/10.3389/fpsyg.2015.00858)
- Hiiippala, T. (2012). Reading paths and visual perception in multimodal research, psychology and brain sciences. *Journal of Pragmatics*, 44(3), 315–327. doi:[10.1016/j.pragma.2011.12.008](https://doi.org/10.1016/j.pragma.2011.12.008)
- Holmqvist, K., Nyström, N., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (Eds.) (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

- Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, 56(3), 373–394. doi:10.1177/0023830913484899
- Hood, W. W., & Wilson, C. S. (2001). The literature of Bibliometrics, Scientometrics, and Informetrics. *Scientometrics*, 52(2), 291–314. doi:10.1023/A:1017919924342
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. doi:10.1177/0267658312461803
- Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. doi:10.1515/iral-2015-0014
- Humphreys, G. F., Mirković, J., & Gennari, S. P. (2016). Similarity-based competition in relative clause production and comprehension. *Journal of Memory and Language*, 89, 200–221. doi:10.1016/j.jml.2015.12.007
- Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, 66(1), 265–284. doi:10.1016/j.jml.2011.09.002
- Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, 98, 1–11. doi:10.1016/j.jml.2017.09.002
- Jackson, C. N., Dussias, P. E., & Hristovia, A. (2012). Using eye-tracking to study the on-line processing of case-marking information among intermediate L2 learners of German. *International Review of Applied Linguistics in Language Teaching*, 50, 101–133.
- Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in educational science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, 10(1), 1–18. doi:10.16910/jemr.10.1.3
- Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year-olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition, and Neuroscience*, 29(7), 877–892. doi:10.1080/01690965.2013.804941
- Julie, F., Millotte, S., Posada, A., & Rizzi, L. (2013). Abstract knowledge of word order by 19 months: An eye-tracking study. *Applied Psycholinguistics*, 34, 323–336. doi:10.1017/S0142716411000713
- Jung, H. (2005). CALL: Past, present and future — a bibliometric approach. *ReCALL*, 17(1), 4–17. doi:10.1017/S0958344005000212
- Kaiser, E. (2011). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes*, 26(10), 1587–1624. doi:10.1080/01690965.2010.522915
- Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113–147. doi:10.1016/j.cognition.2004.01.002
- Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, 57(1), 119–163. doi:10.1111/j.1467-9922.2007.00401.x
- Kim, E., Montrul, S., & Yoon, J. (2015). The in-line processing of binding principles in second language acquisition: Evidence from eye-tracking. *Applied Psycholinguistics*, 36(6), 1317–1374. doi:10.1017/S0142716414000307

- Kleinberg, J. (2002). Bursty and hierarchical structure in streams. *Paper Presented at the Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, Alberta, Canada. Retrieved from <http://www.cs.cornell.edu/home/kleinber/bhs.pdf>.
- Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science*, 30(3), 481–529. doi:10.1207/s15516709cog0000_65
- Kredel, R., Vater, C., Klostermann, A., & Hossner, E.-J. (2017). Eye-tracking technology and the dynamics of natural gaze behavior in sports: A systematic review of 40 years of research. *Frontiers in Psychology*, 8, 1845. doi:10.3389/fpsyg.2017.01845
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. (1st ed.). New York: University of Chicago Press. doi:10.1086/ahr/68.3.700
- Kuperman, V., Bertram, R., & Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, 62(2), 83–97. doi:10.1016/j.jml.2009.10.001
- Kush, D., Lidz, J., & Phillips, C. (2015). Relation-sensitive retrieval: Evidence from bound variable pronouns. *Journal of Memory and Language*, 82, 18–40. doi:10.1016/j.jml.2015.02.003
- Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, and Hearing Research*, 42(3), 526–539. doi:10.1044/jslhr.4203.526
- Lee, J., Yoshida, M., & Thompson, C. K. (2015). Grammatical planning units during real-time sentence production in agrammatic aphasia and healthy speakers. *Journal of Speech, Language, and Hearing Research*, 58, 1182–1194. doi:10.1044/2015_JSLHR-L14-0250
- Lei, L., & Liu, D. (2019). The research trends and contributions of System's publications over the past four decades (1973–2017): A bibliometric analysis. *System*, 80, 1–13. doi:10.1016/j.system.2018.10.003
- Lowder, M. W., & Gordon, P. C. (2012). The pistol that injured the cowboy: Difficulty with inanimate subject–verb integration is reduced by structural separation. *Journal of Memory and Language*, 66(4), 819–832. doi:10.1016/j.jml.2012.03.006
- Mak, W. M., & Sanders, T. J. M. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes*, 28(9), 1414–1437. doi:10.1080/01690965.2012.708423
- Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 57(3), 203–220. doi:10.1037/h0087426
- McMurray, B., Clayards, M. A., Tanenhaus, M. K., & Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin & Review*, 15, 1064–1071. doi:10.3758/PBR.15.6.1064
- Meara, P. (2012). The Bibliometrics of vocabulary acquisition: An exploratory study. *RELC Journal*, 43(1), 7–22. doi:10.1177/0033688212439339
- Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist*, 50(1), 31–42. doi:10.1080/00461520.2015.1004068
- Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypothesis. *Journal of Memory and Language*, 59(3), 266–293. doi:10.1016/j.jml.2008.06.002

- Mitterer, H., & Reinisch, E. (2013). No delays in application of perceptual learning in speech recognition: Evidence from eye-tracking. *Journal of Memory and Language*, 69(4), 527–545. doi:10.1016/j.jml.2013.07.002
- Miwa, K., & Dijkstra, T. (2017). Lexical processes in the recognition of Japanese horizontal and vertical compounds. *Reading and Writing*, 30(4), 791–812. doi:10.1007/s11145-016-9700-6
- Myung, J., Blumstein, S. E., Yee, E., Sedivy, J. C., Thompson-Schill, S. L., & Buxbaum, L. J. (2010). Impaired access to manipulation features in Apraxia: Evidence from eye-tracking and semantic judgment tasks. *Brain and Language*, 112(2), 101–112. bandl.2009.12.003 doi:10.1016/j
- Nalimov, V. V. E., & Mul'chenko, Z. M. (1971). *Measurement of science. Study of the development of science as an information process*. Washington, DC: Foreign Technology Div Wright-Patterson AFB Ohio.
- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and the recognition of printed words. *Brain and Language*, 88, 294–311. doi:10.1016/S0093-934X(03)00168-8
- Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition, and Neuroscience*, 29(2), 186–217. doi:10.1080/01690965.2012.758297
- Puurtinen, M. (2018). Eye on music reading: A methodological review of studies from 1994 to 2017. *Journal of Eye Movement Research*, 11(2), 1–16. doi:10.16910/jemr.11.2.2
- Qiu, J., Zhao, R., Yang, S., & Dong, K. (2017). *Informetrics: Theory, methods and applications*. Singapore: Springer.
- Rader, N. V., & Zukow-Goldring, P. (2012). Caregivers' gestures direct infant attention during early word learning: The importance of dynamic synchrony. *Language Sciences*, 34(5), 559–568. doi:10.1016/j.langsci.2012.03.011
- Ratiu, I., & Azuma, T. (2017). Language control in bilingual adults with and without history of mild traumatic brain injury. *Brain and Language*, 166, 29–39. doi:10.1016/j.bandl.2016.12.004
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. doi:10.1037/0033-2909.124.3.372
- Reali, C., Esaulova, Y., & von Stockhausen, L. (2015). Isolating stereotypical gender in a grammatical gender language: Evidence from eye movements. *Applied Psycholinguistics*, 36(4), 977–1006. doi:10.1017/S0142716414000010
- Reinisch, E., & Sjerps, M. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *Journal of Phonetics*, 41(2), 101–116. doi:10.1016/j.wocn.2013.01.002
- Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, 35(2), 213–235. doi:10.1017/S0272263112000861
- Roland, D., Mauner, G., O'Meara, C., & Yun, H. (2012). Discourse expectations and relative clause processing. *Journal of Memory and Language*, 66(3), 479–508.
- Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, 87, 105–127. doi:10.1016/j.jml.2015.09.005
- Ryskin, R., Qi, Z., Covington, N. V., Duff, M., & Brown-Schmidt, S. (2018). Knowledge and learning of verb biases in amnesia. *Brain and Language*, 180–182, 62–83. doi:10.1016/j.bandl.2018.04.003

- Schmiedtova, B. (2011). Do L2 speakers think in the L1 when speaking in the L2?. *Vigo International Journal of Applied Linguistics*, 8, 139–179.
- Sheppard, S. M., Walenski, M., Love, T., & Shapiro, L. P. (2015). The auditory comprehension of Wh-Questions in aphasia: Support for the intervener hypothesis. *Journal of Speech, Language, and Hearing Research*, 58(3), 781–797. doi:10.1044/2015_JSLHR-L-14-0099
- Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, 58(2), 574–608. doi:10.1016/j.jml.2007.08.001
- Stites, M. C., Federmeier, K. D., & Christianson, K. (2016). Do morphemes matter when reading compound words with transposed letters? Evidence from eyetracking and event-related potentials. *Language, Cognition, and Neuroscience*, 31(10), 1299–1319. doi:10.1080/23273798.2016.1212082
- Suvorov, R. (2015). The use of eye tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing*, 32(4), 463–483. doi:10.1177/0265532214562099
- Tibély, G., Pollner, P., Vicsek, T., & Palla, G. (2013). Extracting tag hierarchies. *PLoS One*, 8(12), e84133. doi:10.1371/journal.pone.0084133
- Tremblay, A., & Spinelli, E. (2013). Segmenting liaison-initial words: The role of predictive dependencies. *Language and Cognitive Processes*, 28(8), 1093–1113. doi:10.1080/01690965.2012.689306
- Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: Evidence from eye-tracking. *Journal of Child Language*, 43(2), 425–441. doi:10.1017/S0305000915000173
- Trude, A. M., Tremblay, A., & Brown-Schmidt, S. (2013). Limitations on adaptation to foreign accents. *Journal of Memory and Language*, 69(3), 349–367. doi:10.1016/j.jml.2013.05.002
- Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes*, 27(7-8), 979–1001. doi:10.1080/01690965.2011.597153
- Underwood, G., Jebbett, L., & Roberts, K. (2004). Inspecting pictures for information to verify a sentence: Eye movements in general encoding and in focused search. *The Quarterly Journal of Experimental Psychology*, 57A, 165–182.
- Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 1–32). Oxford, NY: Oxford University Press.
- von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, 28(10), 1545–1578. doi:10.1080/01690965.2012.728232
- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 75(2), 105–143. doi:10.1016/S0010-0277(00)00063-9
- Wilkinson, K. M., & Light, J. (2011). Preliminary investigation of visual attention to human figures in photographs: Potential considerations for the design of aided AAC visual scene displays. *Journal of Speech, Language, and Hearing Research*, 54(6), 1644–1657. doi:10.1044/1092-4388(2011/10-0098)
- Winke, P., Sydorenko, T., & Gass, S. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, 97(1), 254–275. doi:10.1111/j.1540-4781.2013.01432.x

- Wolter, L., Gorman, K. S., & Tanenhaus, M. K. (2011). Scalar reference, contrast and discourse: Separating effects of linguistic discourse from availability of the referent. *Journal of Memory and Language*, 65(3), 299–317. doi:10.1016/j.jml.2011.04.010
- Yeldham, M. (2018). Viewing L2 captioned videos: What's in it for the listener?. *Computer Assisted Language Learning*, 31(4), 367–389. doi:10.1080/09588221.2017.1406956
- Yi, L., Fan, Y., Quinn, P. C., Feng, C., Huang, D., Li, J., ... Lee, K. (2013). Abnormality in face scanning by children with autism spectrum disorder is limited to the eye region: Evidence from multi-method analyses of eye tracking data. *Journal of Vision*, 13(10), 5. doi:10.1167/13.10.5
- Yoon, S., & Brown-Schmidt, S. (2013). Lexical differentiation in language production and comprehension. *Journal of Memory and Language*, 69(3), 397–416. doi:10.1016/j.jml.2013.05.005
- Yu, P., Pan, Y., Li, C., Zhang, Z., Shi, Q., Chu, W., ... Zhu, Z. (2016). User-centred design for Chinese-oriented spoken English learning system. *Computer Assisted Language Learning*, 29(5), 984–1000. doi:10.1080/09588221.2015.1121877
- Zhang, J., & Kalinowski, J. (2012). Culture and listeners' gaze responses to stuttering. *International Journal of Language & Communication Disorders*, 47, 388–397. doi:10.1111/j.1460-6984.2012.00152.x
- Zhou, P., Su, Y., Crain, S., Gao, L., & Zhan, L. (2012). Children's use of phonological information in ambiguity resolution: A view from Mandarin Chinese. *Journal of Child Language*, 39(4), 687–730. doi:10.1017/S0305000911000249

Appendix

Summary of clusters 5 through 15

Cluster 5: Healthy (neurotypical) bilinguals' control

There is a myriad of eye tracking study designs and methodologies to investigate various research questions in psychology and psycholinguistics. Some common areas of interest involve populations of neurotypical secondary language learners and children and neurodivergent individuals. In cluster #5, Järvikivi, Pyykkönen-Klauck, Schimke, Colonna, and Hemforth (2014) used a visual world eye tracking paradigm with neurotypical 4-year-olds to investigate pronoun resolution, and found that the results were similar to those for a control group of adults. Eye tracking methodologies are also used with individuals with specific conditions. For instance, Lee, Yoshida, and Thompson (2015) and Sheppard, Walenski, Love, and Shapiro (2015) recruited participants with aphasia (agrammatic aphasia and Broca's aphasia, respectively) to investigate their language abilities, while Ratiu and Azuma (2017) examined bilingual individuals with mild traumatic brain injury and their various executive function abilities. In Ratiu and Azuma's study, eye tracking measures allowed finer grain analysis of language impairments in neurodivergent individuals, with fixation patterns providing a wealth of information on language proficiency. In sum, eye tracking can provide

important information on various cognitive abilities and processes in both adults and children from various populations.

Cluster 6: Target language in speech perception

Speech perception in processing various target languages is known to be dependent on a myriad of cues from either the external environment or the speaker (Reinisch & Sjerps, 2013). Winke et al. (2013) study, which was reviewed earlier, showed significant differences between gaze patterns in adults with different language backgrounds, indicating differential language processing. Another study in this cluster is Roberts and Siyanova-Chanturia (2013), which is a review stressing the role of eye tracking in investigating L2 processing. Mitterer and Reinisch (2013) conducted a study of 26 participants who completed a learning phase and a test phase as follows. In the learning phase, one group of participants heard an ambiguous fricative that replaced either the /s/ or /f/ ending sound of a word. In the test phase, participants went through the same procedures, but with fewer trials, and were asked to click on the word they heard after each trial. A perceptual learning effect was later found, whereby participants who were exposed to /f/ fricatives in the learning phase were more likely to fixate on f-final words than their counterparts. The opposite effect was true for individuals who were exposed to /s/ fricative in the learning phase. Hence, this experiment highlights the role that prior learning plays in subsequent speech perception.

Various other characteristics of speech can affect perception and comprehension. Breen and Clifton (2011) found a cost to reading if stress patterns do not conform to expectations, as indicated by a longer fixation duration measures. Canestrelli, Mak, and Sanders (2013) and Mak and Sanders (2013) also found that sentences of causality relationships (i.e., when the first clause provides an explanation of the second clause) were easier to read and comprehend, as indicated by a shorter reading time, than sentences with lower causality. The positive effect of causality on reading spans across claim-argument and consequence-cause sentences (Canestrelli et al., 2013). These results indicate the role that one's own expectations play in speech perception.

Lastly, various speaker characteristics are known to affect speech perception and comprehension, such as the speaker's accent. For differing accents, listeners must deal with variations and make use of differing acoustic information to make sense of the speech (Reinisch & Sjerps, 2013). Trude et al. (2013) found that, while many seem to struggle with foreign accents, one can understand these accents with sufficient exposure, even to near native level, as indicated by faster fixation to the correct target over the course of the experiment. In addition, speakers

provide facial information as they speak, which is an important source of information that listeners attend to. For instance, Yi et al. (2013) found that participants used gaze strategies to shift fixation points towards speakers' mouths to obtain greater comprehension, as indicated by an optimal speech intelligence score. Similarly, Evitts and Gallop (2011) found potential implications of auditory-visual speech perception. In their study, participants viewed the lower facial region of an esophageal speaker, possibly as an additional strategy to cope with the larger cognitive processing load required. Overall, these experiments provide evidence that listeners do consider the facial cues of speakers in an attempt to understand speech.

Cluster 7: Verb bias

In language processing, verbs and sentence structure play central roles in facilitating understanding and comprehension. As a result, the manipulation of verbs during language processing has become a large area of interest. Adding to the current literature, Andreu, Sanz-Torrent, and Trueswell (2013) designed a study to look at how understanding of the argument structure of verbs can guide the selection of subsequent visual information. The study distinguished between three classes of verbs: one-argument, two-argument, or three-argument verbs. Children were presented verb sentences and later asked to look at corresponding images. Eye movements recorded with an Iriscom QuickGlance 2SH eye tracker showed that all children were able to process the verb's semantics and structures to fixate on the correct image, despite the language impairments. The ability to manipulate verbs was also demonstrated in Andreu, Sanz-Torrent, Olmos et al.'s (2013) study comparing adults, children, and neurodivergent children. All three populations showed anticipatory processing of the target as the audio sentences unfolded. In other words, participants were able to use knowledge regarding the semantic properties of verbs to predict the upcoming subject

Extending the research of manipulation of verbs in populations with specific impairments, Ryskin, Qi, Covington, Duff, and Brown-Schmidt (2018) recruited two groups, individuals with amnesia (four with bilateral hippocampal damage and one with severe declarative memory impairments) and neurotypical individuals. In the experiment, participants sat in front of toys and were instructed to carry out actions with the toys (e.g., 'tickle the frog with the feather'). It was found that amnesia patients could still make use of previously acquired verb information to guide processing, as demonstrated by participants' ability to fixate on the correct subject. In sum, the studies in this cluster demonstrate the innate

ability to manipulate verb information to guide sentence and language processing.

Cluster 8: Adult L2 learners

One major field of investigation of L2 learners concerns differences in cognitive differences between L1 and L2 language users/learners. Kim, Montrul, and Yoon (2015) investigated the ability of L2 learners to interpret reflexives and pronouns compared to native speakers. They recruited 44 native English speakers and 32 Koreans with English as L2 and had them listen to auditory instructions containing reflexives or pronouns and then perform the instructed actions. While performing the tasks, participants' eye movements were recorded using an EyeLink 1000. When participants saw the pronouns, L2 learners tended to perform the act incorrectly to a greater extent than native speakers. Such difficulty was not observed under the reflexive condition. For both conditions, the L2 participants took a longer time compared with the other group, as shown by the eye tracking data, indicating failure to bind sentences together and suggesting disparity between L1 and L2 learners. Ellert (2013) investigated the typologically similar German and Dutch languages, and found that L1 and L2 learners applied specific strategies (i.e., different fixation patterns recorded on an EyeLink II) to perform the instructed tasks after listening to mini-stories in either language. Taken together, these studies, in this relatively small cluster, reveal the inherent differences between L1 and L2, despite the level of mastery and how learners employ different strategies and engage in different cognitive processes in L1 and L2.

Cluster 9: Message formulation

Message formulation during language production, although natural to native speakers, requires multiple levels of processing such as retrieval of lexical concepts, building of structures, and subsequent translation of structures to linguistic forms. Brown-Schmidt and Tanenhaus (2006) recognized the dearth of studies examining how utterance planning and message formulation are coordinated, and aimed to address it. A promising methodology to investigate this question is the visual world paradigm, whereby the pattern and timing of eye movements can provide a wealth of information regarding the interplay between message formulation and utterance planning. Participants in Brown-Schmidt and Tanenhaus's study were shown 12 images of various shapes and took turns instructing each other to click on one of the pictures on the screen. It was found that fixations on size contrast predicted the use of size

adjectives. In addition, a high average fixation duration of 600 ms was found before the onset of utterances, demonstrating the complex mechanism of message planning, repairing messages, and production of final utterances.

Building on the literature regarding message formulation and subsequent speech production, Kaushanskaya and Marian (2007) focused on bilinguals and how they exercised control over language selection during production. Theories of language production in bilinguals to date had proposed parallel activation of lexical items for the two different languages. To investigate this theory further, Kaushanskaya and Marian examined how Russian-English bilinguals processed written inputs containing either orthographic or phonological information. The results revealed that Russian participants fixated on non-target Russian inputs that were detected and recognized during an English processing task, reinforcing the idea that both languages were activated in parallel.

The complex nature of language production has rendered this ability susceptible to weakening. For instance, when asked to name objects in a visual scene, as Griffin and Spieler (2006) reported, older speakers took 170 ms longer than younger adults to begin naming. In addition, fixation recorded using an ISCAN 400 remote eye tracker revealed that older individuals spent about 150 ms longer gazing at the object, likely due to the greater retrieval difficulty that comes with slower working memory. However, it is important to note that these findings may be a result of change to vision in older adults. Overall, these studies explored the complexity of formulation during language production that appears to be innate to humans' everyday lives. Such complexities are further demonstrated by how abilities can fade over time, as shown by weakening cognitive abilities.

Cluster 10: Grammatical gender

One prominent topic within L2 acquisition is the establishment of gender agreement. Learning grammatical gender in L2 requires both lexical and syntactic knowledge. Hopp (2013) investigated the ability of L2 learners to acquire new grammatical features and specifically, to process them online in the German language. German has a rich paradigm for gender distinction and is thus a rich source of information. In the study, 20 L2 German speakers and 20 native German speakers participated in an experiment in which they were first shown four still images. Participants were then asked to name the objects and their colors. Later, participants were tested to fixate on the correct image that fit the auditory presentation of sentences. Several interesting results were obtained.

For example, L2 learners used gender on determiners much later than native counterparts, and thus performed weakly at predictive processing.

Grüter, Lew-Williams, and Fernald (2012) designed a more holistic experiment of grammatical gender in L2 learners, employing various measures including offline sentence-picture matching, elicited production, and online processing. During online processing, in which eye movements were recorded, participants viewed pairs of images on the screen while listening to sentences naming the images. In a same-gender trial, the two images depicted objects referred to by nouns of the same grammatical gender. The reverse applied to the different-gender condition. It was found that native speakers had an overall better performance, although the differences between the two groups were small. Even though L2 speakers were of high proficiency, they showed weaker manipulation of gender-marking on determiners that preceded familiar nouns. These prominent studies in the small field of grammatical gender research contribute some valuable findings to the literature.

Cluster 11: Reference resolution

Language comprehension is not an easy feat as it requires an array of skills, such as structure resolution, contrast resolution, and reference resolution. Kaiser (2011) noted the lack of research in some referential forms—specifically strong or emphatic pronouns and the roles they play in interpretation. He thus investigated contrast and salience accounts of Dutch pronouns. The former simply refers to strong vs weak pronouns, while the latter argues that pronouns—whether strong or weak—follow a salience hierarchy (Kaiser, 2011). The more salient the referent, the more prominently it is represented in one's mental model. Kaiser's study of adult Dutch native speakers revealed that reference resolution is a flexible process that incorporates both contrast and salience.

Building on pronoun comprehension, Bergmann, Paulus, and Fikkert (2012) investigated the ability of 3- and 4-years-old children to manipulate pronouns and reflexives. Their study involved children first looking at an image and later hearing auditory material. The children later had to select the image that best corresponds to what they heard with their eye movements recorded using a Tobii system throughout the entire process. Bergmann et al. found that the pronoun and reflexive pronoun condition elicited very different eye movements from participants. Specifically, the younger age group did not fixate differently whether listening to the pronoun or reflexive. Four-year-olds, however, showed a significantly longer fixation during the pronoun condition, possibly indicating the development of knowledge of reflexives when one is at least 4 years old.

Zhou et al. (2012) also looked into ambiguity resolution. Their experiment involved ambiguous Mandarin sentences carrying dual meanings. Participants in this experiment consisted of 44 Mandarin-speaking children with a mean age of 4.5 years. It was found that in the face of ambiguities, participants could correctly comprehend and resolve ambiguities by focusing on stress patterns. The ability for one to focus on other sources of information to resolve ambiguity is further supported by Ito et al. (2012), who found that young children could use pitch prominence to resolve referential ambiguities.

Taken together, these studies demonstrate the usefulness of eye tracking in revealing the ability of various populations to resolve ambiguity or, even more generally, to comprehend texts by incorporating additional information. However, these studies were conducted on a very limited number of languages (e.g., Dutch and Japanese) and, therefore, future studies should look into extending these findings with other languages.

Cluster 14: Interplay account

Various psycholinguistic studies have found support for how linguistic experiences are interlinked with people, events, or the environment (Crocker et al., 2010). To explain this phenomenon, Knoeferle and Crocker (2006) developed the aforementioned CIA, which posits that incremental interpretation of speech drives visual attention towards the anticipated or mentioned objects or events. Such an incremental view of language understanding is supported by various studies. For instance, Corley (2010) investigated gaze fixation as sentences unfolded, especially in the case of language repairs, in 24 university students. Participants were presented with a visual scene and stimulus sentences simultaneously, and their eye movements were tracked using a head-mounted SP Research EyeLink II. Upon hearing the verb, participants were able to fixate on the correct agent in the scene. More importantly, when a repair was made in the sentence, fixations on the agent changed prior to the noun, indicating the constant update of language comprehension as the sentence unfolded.

Similarly, Coventry et al. (2010) showed participants eight sets of images that depicted a first object placed higher/lower or left/right of a second object while simultaneously presenting sentences. Participants were later instructed to judge how appropriate the sentence was at describing the image scene. The results revealed a shifting allocation of attention to objects in the scene as a function of a judgement of how the objects interact as deciphered from the presented sentences. This result indicates a tendency for one to shift attention and fixate on differing stimuli as language processing unfolds.

Farmer, Cargill, Hindy, Dale, and Spivey (2007) brought these findings beyond listening and reading of text by incorporating motor control. In their study, participants were shown specific scenes containing four to six objects and asked to carry out specific actions using the mouse. The study found an eye movement–mouse movement sequence even before the speech file had ended. This result indicates a tendency for switching visual attention and motor control as one processes language incrementally. Overall, these studies demonstrate intertwined relations between language processing, visual attention, and even motor action.

Cluster 15: Anticipatory eye movement

One important benefit of eye tracking methodologies is the ability to measure real-time responses to texts and stimuli. Similar to how research has shown that language is predictive in nature, eye-movement has also been found to be predictive. Verb-based anticipatory abilities of children have been widely established in the literature (Fernald et al., 2006). Andreu, Sanz-Torrent, Olmos, et al. (2013) study, which was previously described, found that argument structures can guide visual attention and sentence production. Along the same lines, Andreu, Sanz-Torrent, and Trueswell (2013) found that different populations showed anticipatory processing of a target as an audio sentence unfolded. Although anticipatory eye gazes were lower in children with specific language impairment compared to adults, they were comparable to matched children in the control group. Taken together, these results demonstrate the innate tendencies for anticipatory eye movements across various populations.