Co-occurrence and Co-authorship Analysis in Vocabulary Acquisition

Ni Putu Meira Purnama Yanti[(✉)] and Mirsa Umiyati

Master of Linguistics Program, Warmadewa University, Denpasar, Indonesia
meirapurnama1504@gmail.com

Abstract. This study is to investigate the current status of annual publications, the keywords of research hotspots, the co-occurrence of keywords, and co-authorship. A total of 1,494 references that were downloaded from Scopus are evaluated. The VOSviewer application is utilized for the analysis. This study presents numerous data on annual trends, the distribution of keywords, the co-occurrence of keywords in 10 clusters, the country and institute co-authorship, the most cited VA-associated publications, and also the reference and journal co-citation. This study examines the state and trends of vocabulary acquisition development. It can assist individuals and academic scholars in gaining a thorough understanding of vocabulary development. It also serves as a reference for the study and implementation of vocabulary acquisition (VA) visualization approaches.

Keywords: *VOSviewer · Vocabulary acquisition · co-citation*

1 Introduction

The increase in research outputs in the last twenty years has given rise to a growing interest in bibliometric analysis. Bibliometric is now being used to evaluate academic outcomes quantitatively. It provides coherent accounts of the trends in a large body of research that could in any other case be intractable. A research trend is the collective action of a group of researchers, each of whom begins to devote substantial attention to a particular scientific topic: they read scientific papers on the subject, cite them, and publish the results of their own research [1]. In any context, research trends can be analyzed. One of these is vocabulary development.

Vocabulary acquisition refers to an essential part of the communication of meaning [2] and of mastering language [3]. True vocabulary development requires more than the capacity to describe words and deduce their basic meanings from context. The ability to discuss, elaborate, and use numerous contexts to demonstrate a clear comprehension of word meanings is necessary for vocabulary acquisition. Vocabulary acquisition is the most in-depth learning process, involving reasoning, perception, memory, and storage of words. It is acquired by reading and is measured by a vocabulary exam that evaluates spelling, part of speech, word meaning, and word production.

Many researchers have conducted research on vocabulary acquisition [4]. investigated about dimensions of vocabulary acquisition, namely labeling, packaging, and

network building. [5] presented that vocabulary acquisition comes from teacher explanation and repeated listening to stories. Vocabulary acquisition is also possible from storytelling [6]. Vocabulary acquisition through multimedia-enhanced computer-assisted word acquisition program, called CAVOCA [7]. The most effective learning tool of vocabulary acquisition is vocabulary notebook [8]. From 1999 to 2009, it can be seen that the trends of research on vocabulary acquisition are viewed from the element and strategy. Even though there are still many interesting things related to vocabulary acquisition that can be studied more deeply.

In this case, bibliometric analysis can be used to assist in identifying the year of publication, the number of publications that have been carried out in that year, the distribution of keywords, the co-occurrence, and also the co-authorship. The results of the identification can be reviewed so that further research trends in vocabulary acquisition can be seen, both those that are often studied and those that are rarely studied. The results of the review of these trends can then be taken into consideration by researchers to conduct the latest research related to vocabulary acquisition, both interdisciplinary and multidisciplinary.

2   Method

The data for this study were extracted from the Scopus database on February 15, 2022. The Scopus database facilitates the finding of research and enables researchers to claim their work, so ensuring proper recognition and promoting career advancement. It provides online access to a variety of materials, including a vast citation database, making it one of the most complete bibliographic sources available. Scopus provides in-depth author and institution profiles that have been manually curated to ensure accuracy and recall.

For the purpose of this study, it is of interest to investigate the trends of research throughout the year and the number of publications in each year, the keywords of research hotspots, the co-authorship, and the co-occurrence of the keywords associated with "vocabulary acquisition" Using "vocabulary acquisition" as the search term in the Scopus database, "article" is defined as the type of literature. The data were then analyzed using scientific mapping.

Science mapping refers to essential procedure of bibliometrics [9]. It can represent the state and development status of the discipline [10]. VOSviewer is one of the many apps for bibliometrics analysis. It is financed by Leiden University's Centre for Science and Technology Studies. In this work, VOSviewer was utilized for visualizing mapping. VOSviewer is an effective data visualization tool. It is used to collect quantitative and visual data in particular sectors. VOSviewer is a free software application created by [11]. It has an important role in co-occurrence.

3   Results and Discussion

In this section, the outcomes of this investigation are discussed in detail. In Sect. 3.1, the present state of vocabulary acquisition research is described. The keywords analysis of the study hotspots on vocabulary acquisition is introduced in Sect. 3.2. In Sects. 3.3 and 3.4, the co-occurrence and co-authorship analyses are displayed.

3.1   The Current Status of Vocabulary Acquisition (VA)

In this section, the annual patterns of vocabulary acquisition-related articles are discussed.

3.1.1   The Annual Trends of VA-Related Publication

The annual trends of VA-related publications are depicted in Fig. 1. Since 1996, when the first paper on VA-related research was published, its growth has been quite sluggish throughout the subsequent two decades. Prior to 2009, and especially after 2011, there were fewer scholars conducting study in this topic. This resulted in an increase in the quantity of publications. There are numerous causes for the quick expansion. First, the rapid development of Internet technology has made it easier for people to get vast amounts of data on language learning. However, there was a decline in vocabulary acquisition publications between 2014 and 2016. This decrease was not too high, only 2,17%. It might happen due to some researchers were not too interested in conducting VA or even they had difficulties in determining scope of discussion and the relevant theoretical basis of VA. Meanwhile, the publications of VA publications have increased significantly again from 2017 to 2021. This increase could occur because more researchers were finally interested in VA after attending seminars/workshops related to VA or even looking for several article publications on internationally reputed journal sites such as Scopus, Web of Science, Science Direct, and others.

3.2   The Keywords Analysis of VA Related Publications

This section explores the content by examining the distribution of keywords: the keywords co-occurrence network map, the top 10 phrases in VA articles, and the keywords density visualization map. Co-occurrence of terms can effectively identify research hotspots in disciplinary fields, hence advancing scientific investigation [1]. In 1,494 documents pertaining to the VA, 88 keywords were identified. VOSviewer was used to construct the VA keyword co-occurrence network (Fig. 2). The size of the nodes and words in Fig. 2 correlates to the node weights. The weight increases as the node and phrase grow in size. The distance between two nodes shows their connection strength. In general, a shorter distance is indicative of a stronger bond. The connection between two
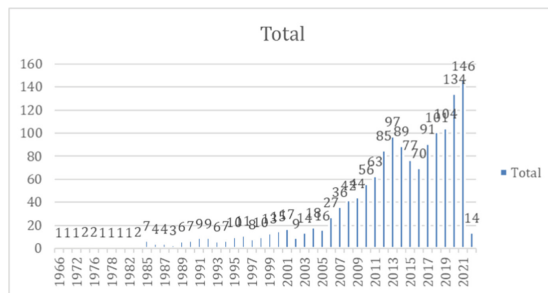

Fig. 1.   The annual trends of VA.related publications.

Fig. 2. Keywords co.occurrence network of VA.related publications

keywords reflects their co-occurrence. The greater the thickness of the line, the more often they appear (Gue et al., 2017). The nodes that share the same color form a cluster. The VA-related literature terms were categorized by VOSviewer into ten separate groups. 330 is the maximum number of times the keyword "vocabulary acquisition" appears. Other commonly appearing terms include "vocabulary" (120 occurrences), "vocabulary learning" (53 occurrences), and "incidental vocabulary acquisition" (13 occurrences) (47).

The link strength between two nodes is their co-occurrence frequency. It can be used as a quantitative index to illustrate the connection between two nodes. (Pinto et al., 2017). The total link strength of a node is the sum of link strengths of this node over all the other nodes. The node, "vocabulary acquisition", has thicker lines with "vocabulary", "reading", "word learning", "vocabulary learning", "incidental vocabulary acquisition", "incidental learning", "language learning", "second language acquisition", and "reading comprehension". The relationship among "vocabulary acquisition", "vocabulary", "reading" imply the components in vocabulary acquisition. The relationship among "vocabulary acquisition", "word learning", "vocabulary learning" represent other names of vocabulary acquisition. The relationship among "vocabulary acquisition", "incidental learning", "language learning", "second language acquisition" infer the domain knowledge of vocabulary acquisition. The association between "vocabulary acquisition" and "reading comprehension" suggests one of the vocabulary acquisition methodologies. Table 1 displays the top 10 keywords together with their frequency and total link strengths.

VOSviewer has the capacity to display density (see Fig. 3). The color of each node in the keywords density visualization graph is determined by the item density at that node. In other words, the hue of a node relies on the number of nearby items. The terms indicated in purple appear more often than those highlighted in yellow. Density views are quite useful for interpreting the overall structure of a map and identifying the most important locations (Chawla et al., 2013). Figure 3 reveals the main research foci of VA study: "vocabulary acquisition," "vocabulary," "vocabulary learning," and "incidental vocabulary acquisition" are significant. These are the essential terms for the VA study.

3.3 The Co-occurrence Analysis on VA

Numerous researchers have using co-occurrence analysis to examine conceptual work in various fields. In this section, the co-occurrence analysis of VA-related publications

Table 1. The top 10 keywords of the VA.related publications

Rank	Keywords	Frequency	Total Link Strength
1	Vocabulary acquisition	330	327
2	Vocabulary	120	131
3	Vocabulary learning	53	57
4	Incidental vocabulary acquisition	47	53
5	Word learning	42	62
6	Reading	42	90
7	Language learning	34	44
8	Second language acquisition	29	38
9	Reading comprehension	22	35
10	Incidental learning	21	47
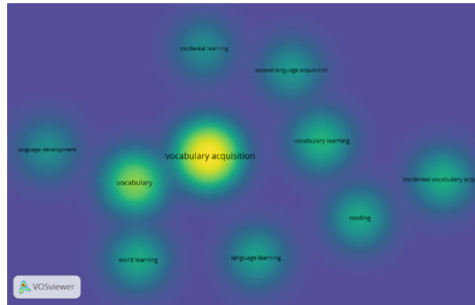
Fig. 3. Keywords density visualization map of VA.related publications

Fig. 4. Co.Occurrence in Three Clusters of VA.related publications

is described, taking into account the relationship between vocabulary learning and the two or three most pertinent terms in three clusters (Fig. 4).

The focus of Cluster 1 is vocabulary acquisition. "second language acquisition" and "incidental learning" are the most applicable terms in this cluster. Second language learning is also known as vocabulary acquisition. Vocabulary acquisition as second language acquisition is the process by which individuals learn vocabularies in a second language

following the acquisition of their native language. The acquisition of L2 vocabulary differs from the acquisition of L1 vocabulary due to the fact that an L2 learner has already constructed a conceptual and semantic framework linked to L1. In addition, one can gain language by incidental learning. Incidental learning refers to the circumstances that allow pupils to acquire vocabulary through incidental means, such as through mental comprehension, and the strategies that can be utilized to encourage and enhance the retention of vocabulary through incidental learning. (Ramos, 2014).

Cluster 2 emphasizes vocabulary. The most appropriate keywords in this cluster are "vocabulary learning", "word learning", and "language learning". Vocabulary is part of vocabulary learning, word learning, and language learning. It is an essential component of language learning. The definitions of new words are usually highlighted. It is the collection of words in a person's language. It improves with age and acts as a basic means of communication and knowledge acquisition. It must be mastered by the student in order to comprehend the language.

Cluster 3 focuses on incidental vocabulary acquisition. "reading" and "reading comprehension" are highly associated inside cluster 3. The process of acquiring several vocabularies through an incidental technique, such as reading, is known as incidental vocabulary acquisition. In order to improve their reading speed and fluency, those who read attentively choose simple, entertaining literature. This is another way of saying that people learn to read through reading, as opposed to analyzing books by studying the vocabulary, grammar, and phrases. Reading comprehension is the ability to read text, process it, and understand its meaning.

### 3.4   The Co-authorship Analysis on VA

It is difficult for a single individual to conduct research on a certain topic. Numerous research endeavors require unquestionably collective expertise. Co-authorship is an integral part of bibliometrics, and the level of research collaboration serves as an indicator of the present state of research in a particular topic (Reyes et al., 2017). This section focuses mostly on nation co-authorship, institute co-authorship, and the most frequently cited works linked to the VA.

3.4.1   The Country Co-authorship Analysis

Co-authorship requires an analysis of nation co-authorship as a fundamental component. It can evaluate both the level of communication between nations and the most influential nations in this field. Figure 5 depicts the co-authorship network for VA-related papers by nation. Large nodes indicate influential nations. The United States, the United Kingdom, Canada, Spain, China, Australia, Germany, the Netherlands, Belgium, and France are, in this regard, the 10 most influential nations.

3.4.2   The Institute Co-authorship Analysis

Figure 6 displays the network of co.authorship between institutes. University of Ottawa from Canada, University of Nottingham from the United Kingdom, Michigan State
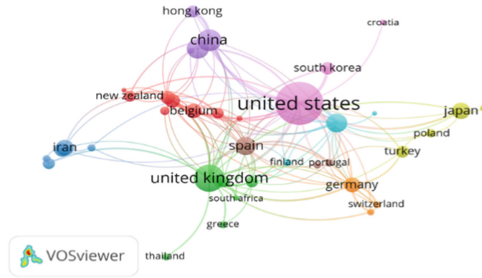
Fig. 5. The country co.authorship network of VA.related publications.
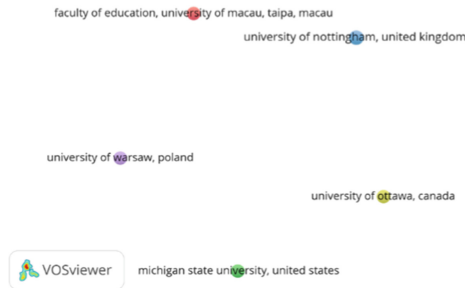
Fig. 6. The institute co.authorship network of VA.related publications.

University from the United States, University of Warsaw from Poland, and University of Macau from Taiwan are the five most influential institutions for VA.related studies.

3.4.2.1 The Most Cited VA Related Publications

To determine the most influential research in VA studies, the ten studies with the most citations are chosen. According to title, journal, authors, publication year, and total citations, Table 2 presents the most frequently cited research.

With 1272 citations, B. Laufer's essay was ranked top. In this study, reading coupled with a word-focused task or a word-focused task without reading proved to be more beneficial for vocabulary improvement. Schmitt N.'s paper ranked second with 1015 citations and discussed the development of a practical model of vocabulary acquisition, vocabulary knowledge from receptive to productive mastery, incorporation of lexical teaching/learning principles into vocabulary and language textbooks, extramural language exposure and how it can most effectively facilitate vocabulary acquisition, informative measures of vocabulary knowledge, and fluency as a component of vocabulary acquisition. The third-ranked paper has 252 citations and is written by E. Peters. Although students who were notified of an upcoming vocabulary exam consulted dictionaries more frequently than those who were not, neither word retention nor text comprehension were affected by test notification. The work authored by Zou d. scored fourth with 166 citations. The author noted that cloze-exercises should be rated as "moderately difficult" because they do not need chunking, hierarchical organization, or pre-task planning. Sentence writing needs chunking and pre-task planning at the sentence level,

Table 2. The Most Frequently Cited Research

Title	Journal	Authors	Year	Citation
Do Second Language Learners Acquire the Majority of Their Vocabulary Through Reading? Some Empirical Evidence	Canadian Modern Language Review	Laufer B.	2003	1272
A research agenda for comprehending vocabulary acquisition, instruction, and evaluation	Language Teaching	Schmitt N.	2019	1015
The effect of task difficulty on L2 vocabulary learning and understanding	Investigating Tasks in Formal Language Learning	Peters E.	2006	252
Extending the evaluation component of the participation load hypothesis to include vocabulary acquisition through cloze activities, sentence writing, and composition writing	Language Teaching Research	Zou d.	2017	166
The game's inbuilt CALL system facilitates the acquisition and pronunciation of English language.	Educational Technology and Society	Young S.S.-C., Wang Y.-H.	2014	140
A diagonalized newton algorithm for non-negative sparse coding	ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings	Van Hamme H.	2013	102
Cognitive psychological method of English vocabulary learning	Revista Argentina de Clinica Psicologica	Xie H., Wu J.	2020	93
To Develop a Data-Driven Learning Model: The Role of Learner Factors in Corpus-Based Second Language Vocabulary Acquisition	Modern Language Journal	Lee H., Warschauer M., Lee J.H..	2020	90

(*continued*)

Table 2. (*continued*)

Title	Journal	Authors	Year	Citation
Which Elements Matter? Building Word Cards for English Vocabulary Development	SAGE Open	Reynolds b. l	2020	83
The usefulness of texting to improve academic vocabulary learning: the perspective of English language learners	Computer Assisted Language Learning	Li J., Cummins J., Deng Q.	2017	60
Can less explanation be more? Enhancing vocabulary through straightforward as opposed to elaborate narrative	First Language	Vaahtoranta E., Suggate S., Jachmann C., Lenhart J., Lenhard W.	2018	45

hence sentence writing evaluations should be given a "strong assessment." Attribute a "very strong evaluation" to composition writing assessment because it demands chunking, hierarchical structure, and pre-task planning at the composition level. With 102 citations, Van Hamme H.'s article was ranked fifth. The study by Xie H. and Wu J. scored sixth with 93 citations. Lee H., M. Warschauer, and J.H. Lee authored a work with 90 citations on August 8. The article by Reynolds b. l. ranked ninth on the list with 83 citations. The paper by Vaahtoranta E., Suggate S., Jachmann C., Lenhart J., and Lenhard W. ranks tenth with forty-five citations, while the study by Li J., Cummins J., and Deng Q. ranks ninth with sixty citations.

3.5   The Co-citation Analysis on VA-Related Publications

Co-citation is defined as the citation of two publications in the same article (Small, 1973). It reveals that literatures are cohesive and that their characteristics, such as who engaged in shouting bouts and who slept with whom, alter in perceivable ways over time, regardless of how one defines and what gave rise to the most significant study. The citation and journal co-citation analysis are displayed below.

3.5.1   The Reference Co-citation Analysis

Reference co-citation analysis is a crucial method for identifying the structure and evolution path of a particular domain. The significance of nodes in the reference co-citation network does not reflect the great amount of citations, but rather highlights the study themes that are strongly correlated with VA-related research. Figure 7 demonstrates that the largest node is Nation, i. s. p. (2001). His paper titled "Learning Vocabulary in Another Language," published by Cambridge University Press, outlined the objectives
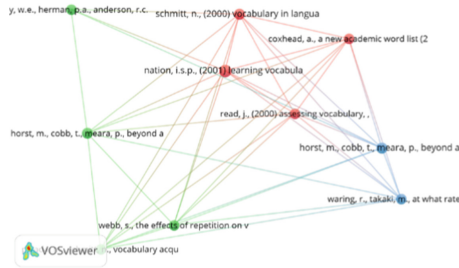
Fig. 7. The reference co.authorship network of VA.related publications
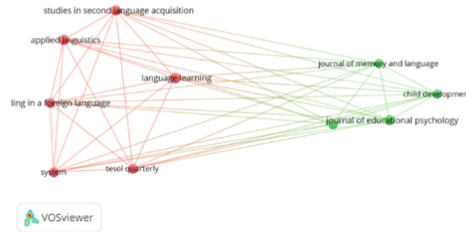

Fig. 8. The journal co.citation network of VA.related publications

of vocabulary learning, specialized uses of vocabulary, vocabulary-learning strategies, learning words from context, word parts, and testing vocabulary knowledge and application. The other significant nodes are Schmitt, N. (2000) and Horst et al. (1998). Schmitt N (2000) in his paper with the title "Vocabulary in Language Teaching" found aspects of knowing a word, incidental and explicit learning of vocabulary, and vocabulary learning strategies. Horst et al. (1998) in their paper entitled "Beyond a clockwork orange: Acquiring second language vocabulary through reading" explained that a small but significant amount of vocabulary learning occurs when reading a simplified novel and having a larger vocabulary size was found to benefit vocabulary acquisition.

3.5.2   The Journal Co-citation Analysis

The examination of journal co-citations is used to determine the broad subject structure and publication characteristics (Hu et al., 2006). VOSviewer is used to visualize the network of journal co-citations for VA-related research. Figure 8 illustrates the top 10 journal co-citation networks from the VA study. Figure 8 demonstrates that the color of each cluster correlates to the group to which it belongs. The network of top ten journal co-citations is separated into two groups. The red cluster includes applied linguistics, language acquisition, reading in a foreign language, second language studies, system, and TESOL quarterly. In contrast, the green cluster includes child development, the journal of educational psychology, the journal of memory and language, and the reading research quarterly.

4 Conclusion

Several intriguing findings regarding VA-related publications are examined and conclusions drawn. Initially, during 1996, VA-related publications fluctuated at a low level. However, following 2017, the number of publications increased fast as more scholars became interested in VA after attending seminars/workshops linked with VA or searching for many articles on internationally renowned journal websites such as Scopus, Web of Science, and others. Second, keyword analysis reveals that vocabulary acquisition is connected with the terms "vocabulary," "vocabulary learning," and "incidental vocabulary acquisition." Thirdly, the co-occurrence of keywords in three clusters of VA-related publications. Cluster 1 is concerned with vocabulary acquisition and includes the keywords "second language acquisition" and "incidental learning." Cluster 2 stresses vocabulary with the following keywords: "vocabulary learning," "word learning," and "language learning." The cluster on incidental vocabulary acquisition that has the highest correlation with the keywords "reading" and "reading comprehension" is Cluster 3. The United States, United Kingdom, Canada, Spain, China, Australia, Germany, Netherlands, Belgium, and France comprise the top 10 country co-authorship networks for VA-related papers. University of Ottawa (Canada), University of Nottingham (United Kingdom), Michigan State University (United States), University of Warsaw (Poland), and University of Macau (Taiwan) are the five most significant institutions among the VA-affiliated publications. Laufer B. with 1272 citations, Schmitt N. with 1015 citations, Peters E. with 252 citations, Zou d. with 166 citations, and Van Hamme H. with 102 citations have the most referenced VA-related papers. Nation, i. s. p., (2001) contains the most co-citations with the article "Learning Vocabulary in Another Language" published by Cambridge University Press. In addition, Schmitt, N. (2000) and Horst et al. (2000) are also commonly co-cited with VA (1998). Schmitt N. (2000) in his article titled "Vocabulary in Language Teaching" and Horst et al. (1998) in their article titled "Beyond a Clockwork Orange: Acquiring Second Language Vocabulary via Reading" discuss the importance of vocabulary in language instruction. VA's top ten journal co-citations are separated into two groups. The red cluster consists of the journals journal of applied linguistics, language acquisition, reading in a foreign language, studies of the second language, system, and TESOL quarterly. The green cluster, meanwhile, consists of the journals journal child development, journal of educational psychology, journal of memory and language, and reading research quarterly.

References

1. Z. Jingxia, Liu., Na, S., & Qing, "Interpersonal Interpretation of Personal Pronoun in Marriage Advertising," *Res. J. English Lang. Lit.*, vol. 3, no. 1, pp. 18–25, 2015.
2. B. Laufer and J. Hulstijn, "Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement," *Appl. Linguist.*, vol. 22, pp. 1–26, 2001.
3. M. Pigada and N. Schmitt, "Vocabulary acquisition from extensive reading: A case study," *Read. a Foreign Lang.*, vol. 18, pp. 1–28, 2006.
4. A. Bryman and R. G. Burgess, *Analyzing Qualitative Data.* London: Routledge, 1999.
5. J. . Moore, Dickson-Deane, and K. Galyen, "E-Learning, Online Learning, and Distance Learning Environments: Are They the Same?," *Internet High. Educ.*, vol. 14, pp. 129–135, 2011.

6. A. Mason, *Demographic transition and demographic dividends in developed and developing Countries.* Mexico: United Nations Expert Group Meeting on Social and Economic implications of Changing Population Age Structure, 2005.
7. A. I. Constantinescu, "Using Technology to Assist in Vocabulary Acquisition and Reading Comprehension," *Internet TESL J.*, 2007.
8. J. Walters and N. Bozkurt, "The effect of keeping vocabulary notebooks on vocabulary acquisition," *SAGE Journals*, vol. 13, no. 4, 2009.
9. H. D. White and K. W. McCain, "Visualizing a discipline: An author co-citation analysis of information science," *J. Am. Soc. Inf. Sci.*, vol. 49, no. 4, pp. 327–355, 1998.
10. V. Garousi and M. V Mäntylä, "Citations, research topics and active countries in software engineering: A bibliometrics study," *Comput. Sci.*, 2006.
11. N. J. Van Eck and L. Waltman, "Software Survey VOSviewer, a Computer Program for Bibliometric Mapping," *Scientometrics*, p. 84, 523–538., 2010.