Complex Dynamic Systems Theory in Second Language Learning and Teaching: A Textometric Review from 2008 to 2022

Huixian Li, Yongyan Zheng*

College of Foreign Languages and Literature, Fudan University, China

Abstract

Complex Dynamic Systems Theory (CDST) has significantly reshaped our understanding of second language learning and teaching across the micro, meso, and macro levels for several decades since its systematic introduction by Larsen-Freeman and Cameron (2008). This study conducts a comprehensive review of empirical research on second language learning and teaching within the framework of CDST from 2008 to 2022. Textometric analysis, a computer-supported qualitative method for textual analysis, was employed to conduct the review by using the IRAMUTEQ software. A total of 198 publications were identified, leading to the emergence of three prominent research themes: research on the language system, research on the second language learner and learning, and research on the second language teacher and teaching. The chronological trends and distinguishing features of the three research themes were discussed against a mainstream model of language learning and teaching at the micro, meso, and macro levels (The Douglas Fir Group, 2016). The findings of this review hold the potential for enlightening future research endeavors in this field.

Keywords: Complex Dynamic Systems Theory, Second Language, Literature Review, Textometry

Introduction

Since Larsen-Freeman introduced the Complex Dynamic Systems Theory (hereafter CDST) into Second Language (L2) research (Larsen-Freeman, 1997; Larsen-Freeman & Cameron, 2008), fundamental questions that have driven research on L2 learning and teaching – how L2 learners deploy their verbal and non-verbal resources at the micro level to engage in specific multilingual contexts at the meso level, while being guided by the overarching ideology towards language use and learning at the macro level (The Douglas Fir Group, 2016) – have undergone paradigm shifts across all levels over the past decades. To gain an understanding of how CDST has reconceptualized these fundamental questions, it is an endeavor at the opportune time, marking the 15th year since its systematic introduction (Larsen-Freeman &...
Cameron, 2008), to synthesize the current CDST-inspired empirical studies, chronicle the developmental trends, and discuss their current status in L2 learning and teaching. For this purpose, the present study reviews the existing CDST-inspired L2 research from 2008 to 2022 to discuss their significance against a mainstream model of language learning and teaching at the micro, meso, and macro levels (The Douglas Fir Group, 2016).

**CDST through an Analytical Lens of a Multilevel L2 Model**

In what follows, we would discuss how CDST provides revolutionary perspectives on L2 learning and teaching across the micro, meso, and macro levels, thus justifying the multilevel model (The Douglas Fir Group, 2016) as an appropriate analytical lens for this review.

At the micro level, CDST views languages and language learners as complex and dynamic systems consisting of multiple components that interconnect and interact with each other to generate some overall states at a specific moment (Larsen-Freeman & Cameron, 2008). Regarding the language system, this perspective rejects the reductionist view of reducing the complex situation to a universal principle that removes all the “noises” in a “grand-sweep” way (Lowie & Verspoor, 2015). Instead, it embraces the belief that language is not fixed or closed but is an ever-developing and open system where patterns emerge from the continuous self-organization among multiple components at multiple scales (Larsen-Freeman, 2015). Rather than being an additive and linear process, language learning is a moment-to-moment effort that constantly feeds into the next condition in a cyclic and iterative fashion (Larsen-Freeman, 2015). There is no target endpoint (Larsen-Freeman, 1997). The growth and decline, acquisition and attrition are equally valuable (de Bot & Larsen-Freeman, 2011). Variability is not merely measurement error but is informative about development (Larsen-Freeman & Cameron, 2008; Verspoor, Lowie, & de Bot, 2021; Verspoor & de Bot, 2021). With regard to the learner system, this perspective corroborates the cognitivism by emphasizing that learners are not only passive recipients of external input. Instead, they actively engage in language learning, and exert a more agentive, robust, and empowered influence over language use and learning by self-organizing their cognitive, affective, interactional, social, political, neural, and semiotic resources (Larsen-Freeman, 2012) and soft-assembling language patterns on given occasions (Thelen & Smith, 1994).

At the meso level, CDST refuses the behaviorism that regards learning and teaching as a uni-directional and linear process from the teacher to the learner (Larsen-Freeman, 2016), or the reductionism which assumes language learning as happening in static isolation or an ideal vacuum with the context as merely a backdrop (Larsen-Freeman & Cameron, 2008). CDST orients for an ecological and holistic perspective recognizing that language learning and teaching are situated in the temporal and spatial contexts which are nested in a hierarchical fashion at different scales (Larsen-Freeman, 2016, 2018). In other words, language is socially constructed (Ellis & Larsen-Freeman, 2006). The language learners, teachers, other stakeholders, and the context are co-adaptive and reciprocally connected systems (Larsen-Freeman & Cameron, 2008), constantly giving and receiving feedback to and from each other (Hiver & Al-Hoorie, 2020). Thus, changes in one system would lead to changes in another (Larsen-Freeman, 2016). Crystalized patterns of behavior or language usage emerge from the co-adaptation of the community at multiple timescales (Larsen-Freeman & Cameron, 2008).

At the macro level, CDST challenges the native/target language standards and the language
teleology – the belief that language learning should aim for an endpoint that meets the native speaker standards (Larsen-Freeman, 2006a). It also denies the linear developmental ladder along which learners climb stage by stage till reach full proficiency (Larsen-Freeman, 2006b). On the contrary, CDST recognizes the existence of variability and fluctuation in the language developmental trajectories, and holds the belief that not all learners are obligated to adhere to the native speaker standards (Larsen-Freeman, 2018), especially those that are shaped by the essentialist language ideology (Ortega, 2017, as cited in Larsen-Freeman, 2018). This also reveals CDST researchers’ rejection of a homogeneous society to embrace a multilingual world, where multiple languages, varieties, and dialects are recognized and their accompanied identities are valued (Larsen-Freeman, 2018). This CDST ideology on language could be reflected in learning and teaching, for example, by designing language assessments that acknowledge learners’ developing capacity rather than their static competence (Larsen-Freeman, 2015), avoiding the “one-size-fits-all” approach to set the same learning goal in the same classroom composed of learners from various linguistic backgrounds (Larsen-Freeman & Tedick, 2016, p. 1339), and even more macro-level language management and planning.

These revolutionary perspectives of CDST have ignited growing research interest in L2 learning and teaching across the micro, meso, and macro levels (see e.g. Fogal, 2022; Hiver, Al-Hoorie, & Evans, 2022), spurring the impetus of this review to synthesize where the current state of CDST-inspired empirical research is and to inform where future research to move. This review distinguishes itself from the previous discussion or systematic/scoping reviews on CDST research (e.g. Fogal, 2022; Hiver, Al-Hoorie, & Evans, 2022; Han, Kang, & Sok, 2023; Larsen-Freeman, 2018) in that it situates the current CDST studies against a more global picture of L2 teaching and learning research at the micro, meso and macro levels, that is, the Douglas Fir Group (2016) model (hereafter the DFG model), in lieu of discussing CDST research per se or in a specific domain. A secondary purpose of this review is to exploit the textometric method as a quantitative tool to complement the traditional qualitative analysis by allowing general themes to emerge in a data-driven and bottom-up fashion. This approach is invulnerable to human bias and capable of foregrounding those underlying patterns that usually go unnoticed by human coding (Ramos, do Rosário Lima, & Amaral-Rosa, 2019). The review is guided by the following research questions:

**RQ1:** What themes emerge from the CDST-inspired L2 research?

**RQ2:** What are the chronological trends of each theme in the CDST-inspired L2 research?

**RQ3:** What is the current status of each theme in the CDST-inspired L2 research?

**Methodology**

**Literature Identification and Screening**

This review follows the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Page, McKenzie, Bossuyt et al., 2021) to identify and screen articles (Figure 1). Firstly, five databases were searched: Web of Science, Scopus, Education Resources Information Center (ERIC), Linguistics and Language Behavior Abstracts (LLBA), and PsycINFO (In’nami & Koizumi, 2010; Oswald & Plonsky, 2010; Plonsky & Brown, 2015), by using keyword combinations: (“foreign language” OR “second language” OR “linguistic” OR “L2”) AND (“Complex Dynamic Systems Theory” OR “complexity theory” OR “dynamic systems” OR “complex dynamic systems” OR “CDST” OR “DST” OR “dynamic
development” OR “complex adaptive systems theory” OR “chaos theory”). Searching was conducted within three research areas: linguistics (language), education, and psychology. The publication time was set from January 2008 to December 2022, as marked by Larsen-Freeman and Cameron’s (2008) seminal work. Papers were restricted to empirical articles published in English, excluding book chapters, theoretical papers, reviews, conference reports, or others. A total of 2761 entries were downloaded from the databases. Then, these identified entries were reviewed by two researchers who are expert in applied linguistics to determine if they meet the criteria of being second language research, and being framed within CDST. By focusing on second language research, we did not intend to make a separation between second language and foreign language, nor did we reject research like first language attrition accompanying L2 development, or the development of languages in addition to the second language. In this review, we used “second language research” as an umbrella term for research on additional languages. Controversial entries were discussed until full agreement was reached (Kappa = 0.82, p < .01). Finally, backward searching has added two paper to the list, ending up with a total of 198 papers (see supplementary materials).

**Figure 1**

*PRISMA Flow Diagram for Identification and Screening Procedures*

- Records identified from databases using keywords and date:
  - Web of Science (n = 1245)
  - Scopus (n = 936)
  - ERIC (n = 115)
  - PsycINFO (n = 44)
  - LLBA (n = 421)
  - Total (n = 2761)

- Records after auto-identification (n = 1110)

- Records excluded (n = 914) by the criteria:
  - Area: second language research
  - Theory: CDST

- Records screened manually by checking the titles, abstracts, keywords, and main texts (n = 196)

- Records removed (n = 1651):
  - Duplicate records in the five databases

- Total studies included in the final pool for reviewing and coding 2008-2022 (n = 198)

- Backward searching for pertinent studies unidentified by the search engine (n = 2)
Textometry as a Method for Literature Review

Textometry is a computer-supported qualitative method for conducting textual analysis (Ramos, do Rosário Lima, & Amaral-Rosa, 2019). It relies on word counts and statistical analysis to conduct data referencing and contextual comparison (Pincemin & Marchand, 2022). In contrast to other textual-analysis methods, textometry effectively maintains a balance between quantitative and qualitative approaches (Pincemin & Marchand, 2022). On the one hand, it shares merits with text mining (Feld & Sanger, 2006), allowing for statistically supported summaries, extractions, and visualizations. On the other hand, it enjoys the advantage of close annotation and thick description, allowing for a rich and detailed view of the text (Pincemin & Marchand, 2022). Compared to the traditional qualitative analysis conducted by human beings, it is not only a robust method capable of handling large volumes of texts with the assistance of quantitative calculation (Pincemin & Heiden, 2008), but also a powerful tool to detect patterns that would otherwise be obscured from naked eyes.

The textometric analysis was carried out using the software Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires (Interface of R for Multidimensional Text and Questionnaire Analysis, hereafter IRAMUTEQ) (http://www.iramuteq.org). It was developed in French in 2009 (Camargo & Justo, 2013) and was later used to conduct discursive textual analysis (Ramos, do Rosário Lima, & Amaral-Rosa, 2019) and textometric review (e.g. Fonseca, Abreu, Guerreiro, & Barros, 2022) in languages other than French. It allows for the bottom-up emergence and visualization of main themes from the corpus texts. The following procedures guided the analysis:

Corpus Building and Cleaning
The abstracts of studies identified from the PRISMA procedure were retrieved to build the textual corpus. The corpus was cleaned by removing the messy information (e.g. citations), checking the spelling of words, ensuring the consistency of expressions and abbreviations, and replacing the unrecognizable punctuation (e.g. quotes, hyphens, percentages, asterisks, etc.) with the underscores recognizable for IRAMUTEQ. In addition, some formulaic academic expressions were made into single forms by connecting the words with underscores so as to avoid the false positivity in the interpretation of word frequency, as exemplified by differentiating the word English in English-as-a-Foreign-Language and Languages-Other-Than-English. Then, we checked and modified the English dictionary in IRAMUTEQ, allowing for a more authentic reflection of the academic texts. For example, the word show was recognized as a noun in IRAMUTEQ’s default English dictionary. But it is more frequently used as a verb in academic abstracts in this review.

Basic Textual Statistics
IRAMUTEQ identifies the number of texts, the frequency of words (types and tokens), hapax legomenon (words occur only once), active (content) words, supplementary (function) words, as well as root-based lemmas (stems). It should be noted that IRAMUTEQ conducts its analysis by segmenting the corpus into units based on text length, typically 40 words per unit by default. These units are considered as the context of words (Lavissière, Sohier, & Lavissière, 2020). In this review, we have chosen to use the length of each abstract (approximately 180 words) as the primary analysis unit, as it better reflects the word context within individual studies.
Descending Hierarchical Classification (DHC) and Factorial Correspondence Analysis (FCA)

The Descending Hierarchical Classification (DHC) allows for detecting the emerging themes in the textual corpus. These themes are derived from word clusters. Specifically, texts are classified into homogenous clusters based on the chi-squared ($\chi^2$) correlation between the cluster and the frequency of the content words (Bart, 2011; Fonseca, Abreu, Guerreiro, & Barros, 2022). Those units containing similar sets of content words would be classified into the same cluster. Thus, different research themes would emerge. This classifying procedure is called the Reinert method (Reinert, 1983). The emerging themes (clusters) are considered as the lexical worlds or semantic contexts (Reinert, 1990), as they represent cognitive “common places” that authors tend to gravitate toward (Reinert, 1993). DHC is distinguished from the traditional qualitative categorization in that the themes are emergent from a data-driven approach and attested by the $\chi^2$ significance, thus avoiding human biases and foregrounding themes that would otherwise be invisible from manual coding (Ramos, do Rosário Lima, & Amaral-Rosa, 2019). The Factorial Correspondence Analysis (FCA) visualizes the distances of each cluster in a Cartesian plane (Ramos, do Rosário Lima, & Amaral-Rosa, 2019).

Data Analysis

To address the first research question, we conducted DHC and FCA on all the abstracts in the corpus to detect the main themes in the 198 CDST research. Analysis was performed on content words except for the verbs (e.g. show, reveal, conduct) for the purpose of obtaining a more accurate picture of the research themes. To address the second research question, first we conducted a manual check and made necessary adjustments to the studies that had been automatically categorized into each theme. Then, we visually depicted the chronological trends of the current research of CDST, including both the overall number of publications and the distribution of publications in each theme. To address the third research question, we made qualitative interpretations on studies in each theme to elucidate their current status in depth.

Findings and Discussion

Themes Emerging from CDST-Inspired L2 Research

Table 1 provides an overview of the descriptive statistics for the abstract corpus. Three clusters emerged from the results of DHC and FCA, successfully classifying 193 out of 198 abstracts (97.47%). In Figure 2, each cluster is visually represented by different colors, and the distances between them are visually depicted. The greater the distance between words or clusters, the more dissimilar they are from each other. The size of words in Figure 2 corresponds to their $\chi^2$ value associated with the respective cluster. Larger word sizes indicate higher $\chi^2$ values, revealing a stronger association between the word and that cluster. Table 2 shows the top 30 most impacted words (words with the highest $\chi^2$) in each cluster.
### Table 1

**Preliminary Analysis after Lemmatization**

<table>
<thead>
<tr>
<th>Index</th>
<th>Total number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of texts</td>
<td>198</td>
</tr>
<tr>
<td>Number of occurrences (tokes)</td>
<td>34207</td>
</tr>
<tr>
<td>Number of lemmas</td>
<td>3260</td>
</tr>
<tr>
<td>Number of forms (types)</td>
<td>4136</td>
</tr>
<tr>
<td>Number of active forms</td>
<td>2344</td>
</tr>
<tr>
<td>Number of supplementary forms</td>
<td>389</td>
</tr>
<tr>
<td>Number of hapax legomenon</td>
<td>1422</td>
</tr>
<tr>
<td></td>
<td>(4.16% of occurrences, 43.62% of forms)</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>3</td>
</tr>
</tbody>
</table>

193 texts classified on 198 (97.47%)

### Figure 2

**DHC and FCA Representation of Impacted Words in Each Cluster**

[Diagram showing clusters 1, 2, and 3 with impacted words categorized by theme.]
Table 2
*The Top 30 Most Impacted Forms in Three Clusters*

<table>
<thead>
<tr>
<th>rank</th>
<th>cluster 1 form</th>
<th>( \chi^2 )</th>
<th>cluster 2 form</th>
<th>( \chi^2 )</th>
<th>cluster 3 form</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>development</td>
<td>34.33**</td>
<td>willingness to communicate to practice</td>
<td>24.11**</td>
<td>32.51**</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>lexical</td>
<td>33.50**</td>
<td>anxiety</td>
<td>21.63**</td>
<td>teacher</td>
<td>24.47**</td>
</tr>
<tr>
<td>3</td>
<td>longitudinal</td>
<td>25.90**</td>
<td>foreign</td>
<td>20.05**</td>
<td>education</td>
<td>20.40**</td>
</tr>
<tr>
<td>4</td>
<td>L2</td>
<td>25.37**</td>
<td>classroom</td>
<td>19.85**</td>
<td>context</td>
<td>17.35**</td>
</tr>
<tr>
<td>5</td>
<td>word</td>
<td>24.06**</td>
<td>enjoyment</td>
<td>18.96**</td>
<td>interview</td>
<td>16.32**</td>
</tr>
<tr>
<td>6</td>
<td>accuracy</td>
<td>23.56**</td>
<td>emotion</td>
<td>16.78**</td>
<td>professional</td>
<td>15.70**</td>
</tr>
<tr>
<td>7</td>
<td>variability</td>
<td>23.43**</td>
<td>idiodynamic</td>
<td>15.13**</td>
<td>educator</td>
<td>15.70**</td>
</tr>
<tr>
<td>8</td>
<td>syntactic</td>
<td>20.40**</td>
<td>factor</td>
<td>14.60**</td>
<td>agency</td>
<td>13.66**</td>
</tr>
<tr>
<td>9</td>
<td>general</td>
<td>18.67**</td>
<td>positive</td>
<td>14.23**</td>
<td>informal</td>
<td>13.66**</td>
</tr>
<tr>
<td>10</td>
<td>fluency</td>
<td>15.98**</td>
<td>communication</td>
<td>11.98**</td>
<td>place</td>
<td>12.22**</td>
</tr>
<tr>
<td>11</td>
<td>regression</td>
<td>15.70**</td>
<td>stable</td>
<td>11.87**</td>
<td>semi-structured</td>
<td>10.96**</td>
</tr>
<tr>
<td>12</td>
<td>sentence</td>
<td>15.70**</td>
<td>interlocutor</td>
<td>11.66**</td>
<td>experience</td>
<td>10.87**</td>
</tr>
<tr>
<td>13</td>
<td>trend</td>
<td>15.70**</td>
<td>variable</td>
<td>11.33**</td>
<td>effective</td>
<td>10.26**</td>
</tr>
<tr>
<td>14</td>
<td>construction</td>
<td>14.22**</td>
<td>participant</td>
<td>11.12**</td>
<td>technology</td>
<td>9.65**</td>
</tr>
<tr>
<td>15</td>
<td>essay</td>
<td>13.66**</td>
<td>turn</td>
<td>10.41**</td>
<td>secondary</td>
<td>9.54**</td>
</tr>
<tr>
<td>16</td>
<td>linear</td>
<td>13.66**</td>
<td>temporal</td>
<td>9.91**</td>
<td>conceptual</td>
<td>9.54**</td>
</tr>
<tr>
<td>17</td>
<td>individual differences</td>
<td>13.35**</td>
<td>class</td>
<td>9.52**</td>
<td>external</td>
<td>9.54**</td>
</tr>
<tr>
<td>18</td>
<td>L1</td>
<td>12.32**</td>
<td>conversation</td>
<td>9.28**</td>
<td>component</td>
<td>8.92**</td>
</tr>
<tr>
<td>19</td>
<td>intra-individual</td>
<td>12.22**</td>
<td>self-ratings</td>
<td>9.28**</td>
<td>educational</td>
<td>8.34**</td>
</tr>
<tr>
<td>20</td>
<td>usage-based</td>
<td>11.65**</td>
<td>self-rated</td>
<td>9.28**</td>
<td>contribution</td>
<td>8.34**</td>
</tr>
<tr>
<td>21</td>
<td>consistent</td>
<td>11.65**</td>
<td>moment-to-moment</td>
<td>9.28**</td>
<td>identity</td>
<td>7.72**</td>
</tr>
<tr>
<td>22</td>
<td>text</td>
<td>11.42**</td>
<td>pedagogical</td>
<td>8.72**</td>
<td>difficulty</td>
<td>7.72**</td>
</tr>
<tr>
<td>23</td>
<td>proficiency</td>
<td>10.87**</td>
<td>recent</td>
<td>8.38**</td>
<td>unique</td>
<td>7.72**</td>
</tr>
<tr>
<td>24</td>
<td>clause</td>
<td>9.65**</td>
<td>nature</td>
<td>8.23**</td>
<td>practical</td>
<td>7.72**</td>
</tr>
<tr>
<td>25</td>
<td>exposure</td>
<td>9.65**</td>
<td>situational</td>
<td>8.12**</td>
<td>COVID 19</td>
<td>7.68**</td>
</tr>
<tr>
<td>26</td>
<td>usage</td>
<td>9.65**</td>
<td>personality</td>
<td>8.12**</td>
<td>awareness</td>
<td>7.68**</td>
</tr>
<tr>
<td>27</td>
<td>path</td>
<td>9.65**</td>
<td>affective</td>
<td>8.08**</td>
<td>author</td>
<td>7.68**</td>
</tr>
<tr>
<td>28</td>
<td>experiment</td>
<td>9.65**</td>
<td>female</td>
<td>7.76**</td>
<td>rich</td>
<td>7.68**</td>
</tr>
<tr>
<td>29</td>
<td>early</td>
<td>9.65**</td>
<td>timescales</td>
<td>7.73**</td>
<td>interpretive</td>
<td>7.68**</td>
</tr>
<tr>
<td>30</td>
<td>Dutch</td>
<td>9.65**</td>
<td>negative</td>
<td>7.73**</td>
<td>conclusion</td>
<td>7.68**</td>
</tr>
</tbody>
</table>

*Note. *\( p < .05, ** p < .01\*

As the results show, Cluster One is related to CDST research on the language systems (e.g. Baba & Nitta, 2014; Polat & Kim, 2014; Spelman & Verspoor, 2010; Verspoor, Lowie & van Dijk, 2008; Yu & Lowie, 2020; Zheng, 2016). The most impacted forms in this cluster include but not limited to development (\( \chi^2 = 34.33, p < 0.01 \)), lexical (\( \chi^2 = 33.50, p < 0.01 \)), longitudinal (\( \chi^2 = 25.90, p < 0.01 \)), word (\( \chi^2 = 24.06, p < 0.01 \)), accuracy (\( \chi^2 = 23.56, p < 0.01 \)), variability (\( \chi^2 = 23.43, p < 0.01 \)), syntactic (\( \chi^2 = 20.40, p < 0.01 \)), fluency (\( \chi^2 = 18.67, p < 0.01 \)), sentence
Huixian Li, Yongyan Zheng

($X^2 = 15.70, p < 0.01$), proficiency ($X^2 = 10.87, p < 0.01$), clause ($X^2 = 9.65, p < 0.01$). For example, Zheng (2016) examined the one-year development of 15 EFL learners’ lexical use and found an increasing trend for lexical diversity and sophistication, a flattening trend for lexical density, and a U-shaped curve for lexical bundles. In another study, Yu and Lowie (2020) examined the oral language development of 10 EFL learners over the course of a semester. Their findings revealed non-linear and dynamic patterns of development, as well as a complex interplay between complexity and accuracy. This interplay evolved from a competitive relationship during the early stages to a more supportive relationship in the later stages of language development.

Cluster Two is related to CDST research on the L2 learner and learning (e.g. Dewaele & Pavelescu, 2019; Han & Hiver, 2018; Kiss & Pack, 2022; MacIntyre & Wang, 2021; Yu, Lowie, & Peng, 2022; Zheng, Lu, & Ren, 2020). The most impacted forms in this cluster include but not limited to willingness to communicate ($X^2 = 24.11, p < 0.01$), anxiety ($X^2 = 21.63, p < 0.01$), enjoyment ($X^2 = 28.96, p < 0.01$), emotion ($X^2 = 16.78, p < 0.01$), affective ($X^2 = 8.08, p < 0.01$). For example, Dewaele and Pavelescu (2019) investigated the changes in two high-school English learners’ willingness to communicate and its relationship with enjoyment and anxiety. They found that various interacting learner-internal and external factors could influence their emotions, which further influences their willingness to communicate in dynamic and idiosyncratic ways. Yu, Lowie and Peng (2022) portrayed 176 EFL learners’ motivational development over two semesters. They observed both a decreasing trend and an increasing trend for learners’ ideal L2 self and ought-to L2 self. They also found three learner profiles: learners with weak ideal L2 self, ought-to L2 self, and learning experience; learners with weak ideal L2 self and ought-to L2 self, but moderate learning experience; and learners with moderate ideal L2 self and learning experience, but weak ought-to L2 self.

Cluster Three is related to CDST research on the L2 teacher and teaching (e.g. Aslan, 2015; Fogal & Koyama, 2022; Smith & King, 2017; Sun & Zhang, 2022; Sak, 2022). The most impacted words in this cluster include but not limited to practice ($X^2 = 32.51, p < 0.01$), teacher ($X^2 = 24.47, p < 0.01$), education ($X^2 = 20.40, p < 0.01$), context ($X^2 = 17.35, p < 0.01$), professional ($X^2 = 15.70, p < 0.01$), educator ($X^2 = 15.70, p < 0.01$), agency ($X^2 = 13.66, p < 0.01$), educational ($X^2 = 8.34, p < 0.01$), identity ($X^2 = 7.72, p < 0.01$), practical ($X^2 = 7.72, p < 0.01$), awareness ($X^2 = 7.68, p < 0.05$). For instance, Sun and Zhang (2022) investigated teachers’ cognition and practice about focus-on-form instruction of the English language by focusing on two novice teachers and two experienced teachers. They found that all teachers favored focus-on-form instruction but the two novice teachers faced complex challenges in implementing focus-on-form instruction in their actual teaching practices. Sak (2022) examined the dynamic changes in two Turkish EFL teachers’ motivation during online classes over two weeks. They found substantial changes within participants and differences across participants in their motivation, originating from a wide range of learner-related, course-related, and personal factors.

The Chronological Trend of each Theme in CDST-Inspired L2 Research
The chronological trends of the number of empirical studies on CDST from 2008 to 2022 is shown in Figure 3. The bar graph represents the number of all publications, which clearly
shows an increasing interest in the area. The first decade (from 2008 to 2017) saw a gradual increase with fluctuation. In the recent five years (from 2018 to 2022), the number of publications surged.

**Figure 3**  
*Chronological Trend of Current Empirical Research on CDST*

![Graph showing the chronological trend of current empirical research on CDST.](image)

Built upon the clusters elicited from DHC and FCA, the colored lines show the chronological trends of each research theme. The blue line represents research on the language systems. This theme of research has arisen researchers’ interest since CDST’s inception, gained popularity between 2013 and 2019 with some fluctuation, and revitalized again since 2020. The green line represents research on the L2 learner and learning. This theme of research increased gradually over the first half period, soared since 2017, and reached a peak in 2022, indicating a growing interest in this theme in recent years. The red line represents research on the L2 teacher and teaching. This theme of research has attracted researchers’ attention quite late, with a majority of studies being published after 2015. This theme is underexplored as compared with the other two themes.

*Current Status of Each Theme in CDST-Inspired L2 Research*

The three themes emerging from the textometric analysis would serve as a framework to guide our qualitative interpretation, anchoring the current status of CDST research in three main areas: the language system, L2 learner and learning, and L2 teacher and teaching. In this section, each of these themes will be discussed in detail and supported by relevant examples.

*Language Systems*

Research on the language systems generally targets at profiling the language developmental
trajectories and detecting the interacting processes of different subsystems (e.g. lexical and clausal subsystems). These studies follow two directions.

One strand of research tracks a small number of cases to elucidate their language developmental paths in depth (e.g. Baba & Nitta, 2014; Polat & Kim, 2014; Spoelman & Verspoor, 2010; Verspoor, Lowie & van Dijk, 2008; Yu & Lowie, 2020; Zheng, 2016), built upon the CDST assumption that language trajectories relying on group average results cannot be generalized to individual trajectories and vice versa (Larsen-Freeman, 2006b). For example, Verspoor, Lowie and van Dijk (2008) observed the writing development of an advanced-level learner over three years. They found that during the first observation point, lexical Type-Token Ratio and sentence length showed a positive correlation but turned into a negative correlation during later points. Spoelman and Verspoor (2010) tracked a Finnish learner over three years with 54 writing samples to examine the accuracy and complexity subsystems. They found the two subsystems developed in non-linear ways, characterized by peaks and regressions, interaction, and competition. Baba and Nitta (2014) traced two EFL university learners over a school year at weekly intervals. Their analysis revealed that both learners experienced at least one phase shift in their writing fluency, as identified through indicators such as sudden leaps, anomalous variance, divergence, and qualitative changes.

A more recent strand of research draws upon learner corpus to crystalize group patterns shared by a small group of homogeneous learners in a bottom-up fashion (e.g. Baba & Nitta, 2021; Gui, Chen, & Verspoor, 2021; Huang, Steinkrauss, & Verspoor, 2021, 2022; Peng, Lowie, & Jager, 2022; Zhang, Zhang, & Zhang, 2022). While exploring intra-individual variability highlights changes and differences, overamplification of it might blur a more global picture of crystalized patterns. Though generalization and prediction are not the targets of CDST (Larsen-Freeman & Cameron, 2008), it does not mean no global patterns could be detected (Molenaar, 2015). In lieu, CDST believes that certain shared patterns would emerge from language use (Larsen-Freeman & Cameron, 2008). For example, Peng, Lowie and Jager (2022) tracked nine Chinese university learners of English over 10 months, and examined their syntactic complexity development. They found five patterns of developmental trajectories regarding the slope (rate), intercept (initial condition), and autocorrelation (dependency on the previous condition) of syntactic complexity, thus unraveling commonalities in learners’ developmental processes that transcended the individual heterogeneity. Huang, Steinkrauss and Verspoor (2021) operationalized the degree of variability in the development of 22 college-level English learners. Their research indicated that as learners diversified their writing strategies, they ultimately made more gains in their writing skills. Their findings provided supportive evidence asserting variability as a signal of improvement at the group level.

L2 Learner and Learning
CDST has brought a revolutionary view on L2 learner and learning that fully recognizes learners’ agency in the learning process (Larsen-Freeman, 2012, 2019). Thereby, a large volume of research has dedicated to investigating language learners as agentive, cognitive, affective, and embodied beings. Such research generally follows two directions.

One branch of studies attempts to untangle the interaction and developmental paths of learner-internal factors in the learning process (e.g. Dewaele & Pavelescu, 2019; Han & Hiver, 2018; Kiss & Pack, 2022; MacIntyre & Legatto, 2011; MacIntyre & Wang, 2021; Yu, Lowie,
For example, MacIntyre and Wang (2021) employed the idiodynamic method to investigate the moment-to-moment changes in three learners’ willingness to communicate. They found that emotions and communicative intentions influenced the underlying dynamic patterns of learners’ willingness to communicate within a communication event. Zheng, Lu and Ren (2020) tracked 15 Chinese university-level students engaged in learning L2 English and L3 Spanish over a period of 1.5 years with a Q methodology. They found two types of changing motivational profiles among learners. One motivational profile was dominated by a translingual and transcultural orientation and developed towards either more constitutive ideal multilingual selves or more language-specific integrative ideal selves. The other profile was dominated by an instrumental orientation and generated diminishing motivational forces.

Another branch probes into how these learner-internal factors impact learners’ divergent language achievements (e.g. Kliesch & Pfenninger, 2021; Li, Dewaele, & Jiang, 2019; Lowie & Verspoor, 2019; Nematizadeh & Wood, 2019; Wood, 2016), grounded in the assumption that even a minor distinction in the initial conditions or in the interplay of internal and external components during the course of their learning could exert a huge impact on their diverse outcomes (Larsen-Freeman & Cameron 2008). For example, Kliesch and Pfenninger (2021) traced 28 older (age 65+) German-speaking learners of Spanish over 7 months. They found a significant moderating effect of education, age, and multilingualism on the L2 proficiency level and the developmental patterns over time, but hardly any relationship between the development of cognition and socioaffect and of L2 performance. Li, Dewaele and Jiang (2019) examined 1,307 Chinese students’ anxiety and enjoyment, their interaction, their effects on learners’ English achievements, and the effect of English achievements on them. They found negative interaction between the two emotions in three groups of learners with different English achievements. Anxiety was negatively associated with proficiency while enjoyment was positively associated with proficiency in nearly all groups except for one low achievement group. They suggested that learners at lower proficiency levels were more likely to experience heightened anxiety and reduced enjoyment in the language learning process.

**L2 Teacher and Teaching**

The theme of L2 teacher and teaching is more related to the context in which language teaching is situated. This review has shown that the prevalent studies have been conducted in the classroom-based instructional context. CDST acknowledges that language learning and teaching are not isolated, static, or linear processes but are instead situated within, emerging from, and dynamically connected to the temporal and spatial environment (Larsen-Freeman, 2018). This perspective shift has ignited research interest in the practice, perception, and challenges of language teacher and teaching.

A salient trend in CDST research over the years has been ascribed to the co-adaptive practice in language teaching (e.g. Fogal & Koyama, 2022; Feryok & Oranje, 2015; Kostoulas, Stelma, Mercer, Cameron, & Dawson, 2018; Rahman & Singh, 2021; Smith & King, 2017; Sun & Zhang, 2022). While the traditional research on teaching “assumes that causation in classrooms operates unilaterally from the teacher to the students” (Bolster, 1983, p.302), a CDST view of teaching attempts to explore the reciprocal effects between students and teachers. For example, Fogal and Koyama (2022) investigated the contextual affordances and
the co-adaptive processes through journaling and journaling feedback in a classroom-based educational practice. Smith and King (2017) observed a postgraduate L2 classroom and examined the effect of teacher’s elicitation types and wait time on student discourse. They found wait-time interactions could lead to non-linear and feedback-sensitive reactions in the classroom discourse system.

Another line of CDST research pays attention to language teachers’ perceptions, as well as teachers’ motivation, beliefs, awareness, self-efficacy, identity, and agency (e.g. Aslan, 2015; Sak, 2022; Sampson, 2016; Sahin & Yildirim, 2016; Yu, Xu, Jiang, & Chan, 2020; Zheng, 2013). The importance of context not only lies in its presence and influence on language teaching and learning, but also in how they are perceived by agents as offering the basis for action (Larsen-Freeman, 2016; van Lier, 2004). As such, teachers’ perceptions might have a huge impact on their teaching practice as teachers are confronted with diverse situations. For example, Aslan (2015) investigated the cognition, identity, and practice of a teacher with dual language identity – both a French native speaker and a German non-native speaker. Their research revealed that the identity as a dual language teacher as well as the early language learning experiences influenced the teachers’ beliefs about teaching contents and processes.

Language teacher and teaching are confronted with manifold challenges (e.g. Qi & Wang, 2022; Ratih, Kurniawan, Nurhidayat, Prayitno, & Buan, 2021; Sulis, Mercer, Mairitsch, Bobic, & Shin, 2021), such as the trend of globalization and the advancement of technology, which are considered as parameters that disturb teachers’ attractor state and pull the teachers into the compeller state, where they self-organize and make adjustments to the new environment (Larsen-Freeman & Cameron, 2008). For instance, Ratih, Kurniawan, Nurhidayat, Prayitno and Buan (2021) investigated the challenges and adjustments confronting pre-service English teachers when taking an international internship. Their findings revealed that these teachers encountered a range of challenges throughout the stages of pre-, while-, and post-teaching when adapting to a new educational system. Qi and Wang (2022) examined the challenges faced by a Chinese-as-a-Second-Language teacher in a blended classroom, composed of both an offline cohort and an online cohort of students simultaneously. They found the teachers’ agency and action changed in complex and dynamic ways when encountering the new teaching context and technology.

**Conclusion and Future Directions**

So far, we have reviewed and discussed L2 empirical research framed within CDST from 2008 to 2022. A total of 198 studies have been reviewed and three research themes have emerged. Our findings revealed that the inception of CDST-inspired L2 research lies in the investigation of the language systems, unraveling the dynamic developmental trajectories, and seeking emergent group patterns. A second line of research probes into the internal factors of language learners, in an attempt to unpack their interconnected and dynamic structures and development, as well as their impacts on diverse language achievements. Attention has also been paid to the practice, perception, and challenges of language teacher and teaching. All of the three themes are indispensable constituents of L2 research under CDST, as they are reciprocally constructed in a complex and dynamic fashion at the micro, meso, and macro levels (Figure 4).

Against a mainstream model of language learning and teaching (The Douglas Fir Group, 2016), this review has found that the current CDST research is mostly confined to the micro-
level examination of the language systems and agent-internal factors under meso-level social engagement contexts. However, the investigation into the macro-level ideology and practice is still in dearth. Therefore, over and above a synthesis of the current research, we suggest three trends that might inspire future efforts.

**Figure 4**
*Current State of L2 Research Framed within CDST (Adapted from The Douglas Fir Group [2016])*
At the micro level, complex as the language system is, the present review has found that the prevalent CDST research has predominantly concentrated on the lexical and syntactic subsystems, with only a small number of studies on other subsystems (e.g. de Leeuw, Mennen, and Scobbie’s [2013] exploration on the phonetic subsystem). A future attempt is to explore how different subsystems of language (e.g. morphology, phonetics, semantics, pragmatics) would interact and develop at different rates and scales. In addition, our findings have showed that while much of the current attention has been paid to the second language, relatively little has been paid to other languages, with only a few exceptions (e.g. Opitz’s [2013] exploration on L1 attrition, Huang, Steinkrauss and Verspoor’s [2022] investigation on L3 influence). The examination of the bidirectional positive and negative influence among multiple language systems would also be inspiring for future research. On the other hand, despite a branch of the latest CDST studies exploring the group trends in language development, evidence still comes short to confirm CDST claims beyond the individual cases (Bulté & Housen, 2020). The inclusion of more group-level research would be valuable in establishing a robust foundation for supporting CDST claims.

At the meso level, a majority of reviewed CDST studies have been based on the traditional classroom context. On the one side, this has left a large picture of other contexts where language use takes place untouched, e.g. family, neighborhood, workplace, worship place, and social organizations (The Douglas Fir Group, 2016). On the other side, the rapid advancement in technology is poised to reshape our world, compressing time and space (Larsen-Freeman, 2018), and even creating entirely new time and space. Future research would move beyond treating technology only as assisting tools for language learning and teaching, and would instead focus on the new technology-created time and space. This could involve investigating how emerging platforms like online meeting rooms, metaverse environments, and chatbots construct new cyberspaces as settings for language learning and teaching. While this endeavor has already begun in language learning and teaching research beyond the scope of CDST, it is important to note that CDST research has made limited inroads into this area. CDST holds the potential to explore how these technologies offer affordances to learners and teachers, reshaping the way they interact and co-adapt in these evolving digital environments.

At the macro level, CDST embraces a multilingual world, recognizing and valuing the coexistence of multiple languages, varieties, dialects, and their accompanying identities (Larsen-Freeman, 2018). However, despite the much theoretical advancement made by CDST, empirical research concerning the practical implementation of the macro-level CDST ideology of language, language learning and teaching remains largely underexplored. For instance, future efforts could be put into designing language assessments that acknowledge learners’ developing capacity rather than their static competence (Larsen-Freeman, 2015), avoiding the “one-size-fits-all” approach to set the same learning goal in the same classroom composed of learners from various linguistic backgrounds (Larsen-Freeman & Tedick, 2016, p. 1339), and even more macro-level language management and planning.

Finally, this review also validates the feasibility and effectiveness of textometry as a computer-supported qualitative method for conducting literature review. Textometric tools like IRAMUTEQ bear the merit of detecting themes from the big data in a bottom-up way. They bring to light new perspectives, interpretations, and relationships that might otherwise remain
hidden in manual coding (Ramos, do Rosário Lima, & Amaral-Rosa, 2019). Furthermore, they are more efficient and objective as compared with human judgment.

**ORCID**
https://orcid.org/0000-0001-8411-5646
https://orcid.org/0000-0002-0664-3942

**Acknowledgements**
This work has been supported by the Youth Innovative Team on Humanities and Social Sciences of Fudan University awarded to Prof. Yongyan Zheng at Fudan University. Special thanks go to Xuefen Zhang, who collaborated with the first author to screen the reviewed articles.

**Funding**
Not applicable.

**Ethics Declarations**
No, there are no conflicting interests.

**Rights and Permissions**
Open Access
This article is licensed under a Creative Commons Attribution 4.0 International License, which grants permission to use, share, adapt, distribute and reproduce in any medium or format provided that proper credit is given to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if any changes were made.

**References**


**Supplementary Materials**


Huixian Li, Yongyan Zheng


Huixian Li, Yongyan Zheng


Huixian Li, Yongyan Zheng


Pfenninger, S. E. (2021). About the inter and the intra in age-related research: Evidence from a longitudinal CLIL study with dense time serial measurements. *Linguistics Vanguard, 7*(s2), 20200028. [https://doi.org/10.1515/irali-2021-0025](https://doi.org/10.1515/irali-2021-0025)


Pham, G. T. (2016). Pathways for learning two languages: Lexical and grammatical associations within and across languages in sequential bilingual children. *Bilingualism, 19*(5), 928-938. [https://doi.org/10.1017/S1366728915000899](https://doi.org/10.1017/S1366728915000899)


Rosmawati (2014). Dynamic development of complexity and accuracy: A case study in second language academic writing. *Australian Review of Applied Linguistics, 37*, 75-100. [https://doi.org/10.1075/aral.37.2.01ros](https://doi.org/10.1075/aral.37.2.01ros)


Sampson, R. J. (2015). Tracing motivational emergence in a classroom language learning project. *System, 50*, 10-20. [https://doi.org/10.1016/j.system.2015.03.001](https://doi.org/10.1016/j.system.2015.03.001)


