

# Network analysis for modeling complex systems in SLA research

Lani Freeborn\* , Sible Andringa , Gabriela Lunansky and Judith Rispens

University of Amsterdam, Amsterdam, The Netherlands

\*Corresponding author. E-mail: [l.j.v.freeborn@uva.nl](mailto:l.j.v.freeborn@uva.nl)

(Received 07 October 2021; Revised 29 July 2022; Accepted 15 August 2022)

## Abstract

Network analysis is a method used to explore the structural relationships between people or organizations, and more recently between psychological constructs. Network analysis is a novel technique that can be used to model psychological constructs that influence language learning as complex systems, with longitudinal data, or cross-sectional data. The majority of complex dynamic systems theory (CDST) research in the field of second language acquisition (SLA) to date has been time-intensive, with a focus on analyzing intraindividual variation with dense longitudinal data collection. The question of how to model systems from a structural perspective using relation-intensive methods is an underexplored dimension of CDST research in applied linguistics. To expand our research agenda, we highlight the potential that psychological networks have for studying individual differences in language learning. We provide two empirical examples of network models using cross-sectional datasets that are publicly available online. We believe that this methodology can complement time-intensive approaches and that it has the potential to contribute to the development of new dimensions of CDST research in applied linguistics.

## Introduction

In the field of second language acquisition (SLA), complex dynamic systems theory (CDST) is a theoretical paradigm used to study the complex and dynamic nature of language, language use, and language development (Hulstijn, 2020). A complex system is formed out of interactions between multiple internal and external system components. For example, if conceptualizing language development as a complex system, changes in development are dependent on interactions between a learner's internal resources like working memory, motivation, and personality, as well as external, environmental resources like the teacher, learning materials, and language use (van Geert, 1991). These internal and external resources are interrelated, whereby altering one component could in turn alter other components of the system (de Bot et al., 2007). In this way, a complex system is characterized by complete interconnectedness and mutual causality (Larsen-Freeman & Cameron, 2008). Complex systems are inherently

dynamic; systems emerge over time through processes of self-organization and coadaptation between micro- and macro-level system components (Larsen-Freeman, 1997). This means that complex systems are soft-assembled, whereby systems are “more than the sum of their parts, reflecting a multiplicative combination of attributes, experiences and situational factors” (American Psychological Association, 2017). CDST researchers in SLA have acknowledged the impossibility of fully “knowing” a system, as complex systems are characterized by unpredictability and nonlinearity, where changes in the system can be disproportionate to the cause (Larsen-Freeman, 1997). Although a complex system is, by definition, constantly in flux, the system can also demonstrate periods of temporary stability. This is referred to as an attractor state; a self-sustaining state in which interactions “are actively reproduced over time” (van Geert, 2019, p. 168). An attractor state represents higher-order patterns of self-organization within state space, from which the system moves toward or away from over time (Hiver, 2014). To illustrate, an attractor state could refer to the tendency for learners not to participate in class and remain silent (Hiver, 2014).

With the growing recognition that CDST approximates the reality of language development (Hiver & Al-Hoorie, 2020a), more SLA researchers are adopting this framework. However, there are many methodological considerations for conducting empirical research within a CDST paradigm. Some of these include how to operationalize the system, how to assess the influence of contextual factors on the system, as well as macro- and micro-structure considerations (Hiver & Al-Hoorie, 2016). Given the inherent complexities of analyzing dynamic cause-effect relationships between systems and their components, there has been much discussion about suitable methodologies and suggestions of how to enhance our CDST toolbox (de Bot, 2011; Hiver & Al-Hoorie, 2016, 2020a; Hiver et al., 2022).

Hilpert and Marchand (2018) distinguish between three conceptual perspectives to studying complex systems and their accompanying research designs: time-intensive, relation-intensive, and time-relation intensive approaches. Firstly, time-intensive approaches “are used to make inferences about system behavior using closely spaced observations over time” using longitudinal data (Hilpert & Marchand, 2018, p. 192). The second approach, relation-intensive, focuses on identifying the structure of the relationships among individuals or variables in a system using cross-sectional data. Combining the first two approaches, time-relation intensive approaches “are used to make inferences about system behavior using closely spaced, simultaneously collected observations of both within-element change and changing between element relationships” (Hilpert & Marchand, 2018).

The majority of CDST studies in the field of SLA to date have taken time-intensive approaches, typically consisting of case studies characterized by dense data collection with qualitative and descriptive data analyses (Hiver et al., 2022). For the last 30 years, CDST researchers have focused on individual variability and the dynamics of processes (van Geert & van Dijk, 2021). This is not surprising, given that CDST is essentially a theory of change, concerned with how one state develops into another state over time. However, as Hilpert and Marchand (2018) have pointed out, complex systems can be studied from multiple perspectives. Besides analyzing change over time, identifying the structure of a system is also a key aspect of CDST. Expanding our line of inquiry to include relation-intensive approaches could contribute to the development of new dimensions of CDST research and complement time-intensive approaches. While researchers have a diverse selection of methods available for time-intensive approaches, our methodological toolbox for relation-intensive methods is lacking.

In this article, we highlight network analysis as a potential methodology to model complex systems from a relation-intensive perspective. While network analysis can also be used for time- and time-relation intensive approaches, this article is focused on network analysis for relation-intensive approaches only, due to the relative lack of attention that this dimension has received by SLA researchers working within a CDST paradigm. More specifically, we concentrate on *psychological networks*, as opposed to *social networks*. SLA researchers have already explored social network analysis as a suitable research methodology for CDST, for example to model relationships between learners in a classroom and teacher networks as complex systems (Hiver & Al-Hoorie, 2016; Hiver & Al-Hoorie, 2020a; Mercer, 2014). SLA researchers have not yet explored the potential of psychological networks to model psychological constructs that influence language learning as complex systems. The network approach to psychopathology has been used to reevaluate theories of mental disorders (Borsboom *et al.*, 2017; Borsboom & Cramer, 2013) and constructs such as intelligence and cognitive development from a CDST perspective (Kievit, 2020; van der Maas *et al.*, 2006, 2017). In this article we discuss how, similarly to psychology research, individual differences in language learning can be modeled as nomological networks, expanding our relation-intensive methods to include the study of phenomenological constructs. We begin with a brief review of CDST research designs used in the field of SLA to date, in relation to the three different conceptual approaches to studying complex systems as described by Hilpert and Marchand (2018). We then expand discussion on relation-intensive approaches, the least researched dimension in CDST. The remainder of the article discusses potential applications of network analysis. To further aid discussion, we provide two examples of network models that are estimated from publicly available data.

## Research designs in CDST

### *Time-intensive methods*

Most CDST research in applied linguistics is time-intensive, with longitudinal data collection of a single variable (or multiple variables for a single case/participant) to observe micro-level changes in the system over time (Hiver *et al.*, 2022; Hiver & Larsen-Freeman, 2019). Time-intensive studies tend to have dense data collection and small sample sizes, with 40% of studies including a sample size of 10 participants or fewer (Hiver *et al.*, 2022). A particularly researched area is the development of L2 writing over time using measures of complexity, fluency, and accuracy (CAF) (Evans & Larsen-Freeman, 2020; Larsen-Freeman, 2006; Lowie *et al.*, 2017; Lowie & Verspoor, 2019). Some common CDST techniques used in these studies include assessing the degree of variability in developmental trajectories and plotting longitudinal data on min-max graphs for visual inspection. Several studies have used a time-series design based on the view that frequent-enough measurements may be able to capture underlying developmental processes (Van Geert & Steenbeek, 2005). For example, Waninge *et al.* (2014) micro-mapped the motivational dynamics of four students during class time, taking measurements at 5-minute intervals.

Another popular methodology for observing language development is retrodictive modeling (Chan & Zhang, 2021; Evans & Larsen-Freeman, 2020; Nitta & Baba, 2018), based on the idea that because what we observe has already changed, change can be described retrospectively (Larsen-Freeman & Cameron, 2008). Retrodictive methods such as process tracing have been used to study the development of language as well as

individual differences over time. For example, Papi and Hiver (2020) used process tracing of retrospective interviews to examine changes in six learners' motivational principles and Amerstorfer (2020) used process tracing with a combination of classroom observations and retrodictive interviews to explore five learners' strategic L2 development. Some time-intensive CDST studies have also used the "idiodynamic method," a mixed-methods approach to studying affective and cognitive states (MacIntyre, 2012). These time-intensive approaches have provided insights into non-linear L2 developmental processes and intraindividual variation over time.

### *Relation-intensive methods*

In comparison to the number of studies that have taken a time-intensive approach, far fewer CDST studies have taken a relation-intensive approach, which involves exploring the structure of relationships between people or variables within a system with cross-sectional data. As previously mentioned, SLA researchers have noted how social network analysis is a suitable methodology for CDST, for example to analyze relationships between learners in a classroom, teacher networks, or school networks (Mercer, 2014). However, this discussion has been mostly theoretical, with very few empirical studies using social network analysis from a CDST perspective. For example, although some applied linguistics researchers have used social network analysis to map the distribution of conversational topics of bilinguals in different contexts (Tiv et al., 2020) and to assess the impact of social networks in study abroad contexts (Gautier, 2019; Paradowski et al., 2021; Zappa-Hollman & Duff, 2014), these studies are not typically informed by CDST.

While relation-intensive approaches can focus on person-to-person interactions, they can also be used to analyze relations among psychological variables (Marchand & Hilpert, 2018). Taking a variable-centered relation-intensive approach necessitates researchers to engage with psychological constructs on a phenomenological level, and to carefully consider whether their methodology can effectively model complex patterns of relationships among variables. Some SLA researchers have discussed how psychological constructs such as "the self" and L2 motivation can be conceptualized as complex systems (Henry, 2014, 2017; Mercer, 2011a). In one study, Mercer (2011a) took a relation-intensive approach to explore how the self-construct could be conceived of as a complex system. Using qualitative data of a single case study, Mercer (2011a) created a three-dimensional network-based model of a student's self-concepts that she felt to be the most "phenomenologically-real" representation of the data.

Besides this, few SLA researchers have attempted to model psychological constructs as complex systems. There are a handful of CDST studies that are reminiscent of relation-intensive approaches, which used quantitative methodologies often deemed ill-suited for CDST. For example, conceptualizing L2 speech as a complex system, Saito et al. (2020) investigated the effects that 30 different internal and external individual differences had on the pronunciation of 110 L2 English speakers. Due to the large number of variables included in their study, Saito et al. (2020) first conducted factor analysis and then did regression analysis with the extracted factor scores on speech ratings. In another study, Li et al. (2020) positioned themselves within a CDST framework to explore the relationships between individual difference constructs including foreign language classroom anxiety, foreign language enjoyment, self-perceived achievement, and actual English achievement. To analyze data, Li et al. (2020) conducted Pearson correlations to assess relationships between variables and

used multiple regression analysis to assess the combined effect of anxiety and enjoyment on language achievement. While these two studies are fine cross-sectional studies in their own right, CDST scholars have argued that methods such as zero-order correlations and linear regression oversimplify the complex realities of how individual differences influence second language development and have questioned the use of cross-sectional datasets in CDST research (Al-Hoorie & Hiver, 2022; Hiver, 2014). Overall, very few SLA researchers to date have used relation-intensive approaches within a CDST paradigm. There are also seemingly fewer methodologies available for SLA researchers to explore relation-intensive approaches, with more conceptual discussion than empirical studies.

### *Time-relation intensive methods*

Hilpert and Marchand (2018, p. 192) describe time-relation intensive research designs as having “closely spaced, simultaneously collected observations of both within-element change and changing between element relationships.” Only a few SLA studies have analyzed interactions between variables *and* how these interactions change over time. However, these studies cannot be strictly classified as time-relation intensive approaches, as their data collection consisted of only a few time points. For example, Serafini (2017) conducted longitudinal case studies to explore interactions between cognitive and motivational individual differences at varying proficiency levels. Data was collected twice from 87 university students learning L2 Spanish, at the beginning and end of an academic semester. Serafini used Pearson correlations to analyze associations between individual differences at each time point and created scatterplots with regression and Loess lines to visualize relationships between variables and compare differences across proficiency levels. Results showed that the relationship between cognitive abilities and motivational constructs varied at each time point and across learner proficiency levels, indicating that cognitive and motivational subsystems are interdependent. In another study, Piniel and Csizér (2014) investigated changes in 21 students’ motivation, anxiety, and self-efficacy at six time points throughout an academic writing course. To analyze data, Piniel and Csizér used latent growth curve modeling (LGCM) and cluster analysis to group together learners with similar trajectories. Interactions between variables were also analyzed by comparing Pearson correlations between IDs at each time point. Overall, results indicated that language learning experience, ought-to L2 self, and writing anxiety showed a significant level of nonlinear change over time. There was also a strong interrelationship between motivation and anxiety, whereby more highly motivated learners had lower levels of language learning anxiety.

Pfenninger and colleagues (Kleisch & Pfenninger, 2021; Pfenninger, 2020) have also recently explored the use of generalized additive mixed modeling (GAMM) for a time-relation intensive approach to SLA microdevelopment. GAMM is a type of analysis used for time-series data that can consider nonlinear development, iterative processes, and interdependency between variables (Pfenninger, 2020). Pfenninger (2020) used GAMM to analyze the L2 developmental trajectories of four groups of children (N = 91) on different content and language integrated learning (CLIL) programs. The children completed various language tasks four times a year for up to 8 years. Pfenninger also combined GAMM with qualitative data to help identify what contributed to developmental trajectories. Results showed that children had similar L2 trajectories regardless of their age of onset, and that L2 growth was determined by various different

external and internal states across time. In another study, Kliesch and Pfenninger (2021) used GAMM to examine the L2 developmental trajectories of 28 adults (age 64+) on a 7-month beginner's Spanish course. Data was collected each week over 30–32 weeks, which included seven L2 measures, eight cognitive tasks, and measures of well-being and motivation. GAMM revealed both linear and nonlinear increases in L2 proficiency over time, with considerable between-subject variability. While only a few CDST studies have used time-relation intensive methods, findings indicate a complex interplay between external and internal learner differences, which in turn interact with language development in a nonlinear way over time.

### Expanding our research agenda

CDST studies that have incorporated a relation-intensive element to their research design are far less common compared to the number of studies that have taken time-intensive approaches. Despite the fact that “complexity theorists are interested in understanding *the relations* [emphasis in original] that connect the components of a complex system” (Hiver & Larsen-Freeman, 2019, p. 287), to date there have been very few attempts to empirically model these relations. One potential reason behind this is relates to methodological challenges and the view that cross-sectional data, zero-order correlations and linear regression are ill-suited to studying complex systems (Al-Hoorie & Hiver, 2022). Another reason relates to the theoretical challenges of conceptualizing abstract psychological constructs as complex systems. A number of individual differences constructs in language learning have been conceptualized as complex systems, such as motivation (Papi & Hiver, 2020), strategy development (Amerstorfer, 2020), anxiety (Gregersen, 2020), working memory (Jackson, 2020), and willingness to communicate (MacIntyre, 2020). To examine these constructs from a relation-intensive perspective, for example to model L2 motivation as a complex system, we must consider the components that form the system, and how these components align with our measurement instruments. Researchers must also confront “the boundary problem” (Larsen-Freeman, 2017), accepting the theoretical impossibility of measuring a complex system in its entirety, whereby “the whole is greater than the sum of its parts” (Han, 2019, p. 156). Consideration should also be given to the phenomenological validity of equating conceptual and theoretical concepts as systems, and the practical implications this has for a chosen methodology (Hiver & Al-Hoorie, 2016). Mercer (2011b, p. 59) discusses these issues in relation her network-based model of the self-concept, acknowledging the theoretical and empirical difficulty of distinguishing the blurred boundaries between different self-constructs. Despite the challenges of exploring psychological constructs related to language learning from a relation-intensive approach, and the inevitable reductionism this entails, focusing on system structures can offer a perspective that is currently missing from CDST research in SLA.

Take the construct of L2 motivation, for example, which has been much discussed in CDST research (e.g., Dörnyei, 2017; Dörnyei et al., 2015; Henry, 2014, 2017; Hiver & Papi, 2019; Hiver & Larsen-Freeman, 2019; Papi & Hiver, 2020). Most CDST research on L2 motivation has been time-intensive with a focus on observing micro-level changes in a small number of variables over time. Very few CDST researchers have explored L2 motivation from a relation-intensive perspective, although there has been some theoretical discussion of how to conceptualize the structural relationships between motivational constructs as complex systems (Henry, 2014, 2017). The L2 Motivational Self System (L2MSS) is a theoretical paradigm that was developed by

Dörnyei (2005, 2009) that conceptualizes L2 motivation from a self-perspective. The L2MSS is comprised of three phenomenologically constructed concepts, each theorized to be a primary source of motivation to learn an L2: the Ideal L2 Self, the Ought-to L2 Self, and L2 Learning Experience (Dörnyei 2005, 2009). Although the L2MSS was not originally conceptualized as a complex system, it has been conceptually extended to a CDST paradigm (Henry, 2014, 2017). For example, Henry (2017, p. 551) has described the self-concept as a multifaceted dynamic structure, which can be understood as “the product of constant interactions between different subsystems (such as, e.g., self-efficacy and self-esteem).”

Taking a relation-intensive approach to L2 motivation could provide insight into the structural relationships between components of the L2 motivational system, and if this were expanded to a time-relation intensive approach, could potentially identify attractor states. SLA researchers have speculated about how the L2 self-system, in particular the Ideal L2 self, can manifest as an attractor state (Henry, 2017; Hiver, 2014; Waninge et al., 2014), whereby “changes in the vision of the Ideal L2 Self and changes in the distance between it and the actual self, can be conceptualized as changes in attractor state *geometries*” (Henry, 2014, p. 87, emphasis in original). Although longitudinal data is needed to show system self-organization and the emergence of attractor states, cross-sectional data can provide a perspective that is currently missing from CDST research in SLA. As Mercer (2011a) reflects in relation to her network-based model of the self-concept:

Whilst the model out of necessity can only represent a snapshot of a fragment of an individual’s self-concept network in a specific context at a particular time, the essence of the underlying form can be used to fundamentally understand the structure and nature of self-concept. (p. 66)

Taking a relation-intensive CDST approach to the study of individual differences in SLA can thus be viewed as complimentary of time-intensive approaches. Cross-sectional data can provide insight into the structure of relationships between system components, which, if combined with what we have learned from time-intensive CDST studies, could enrich our understanding of the complex interplay between individual differences and L2 development.

There is currently little guidance on how to analyze and model interactions between system components from a relation-intensive perspective. As previously mentioned, CDST researchers have questioned whether methods such as zero-order correlations and linear regression are suitable for examining dynamic changes and interconnected (Al-Hoorie & Hiver, 2022). Although scholars have emphasized the potential of quantitative analyses for CDST research (Al-Hoorie & Hiver, 2022) for example to identify network structure or nested phenomena, there appears to be an overall reluctance to use cross-sectional data, with most CDST researchers preferring longitudinal data. Until now, most studies that have taken a relation-intensive approach have analyzed relationships between variables by correlations and multiple regression analysis (Li et al., 2020; Piniel & Csizér, 2014; Saito et al., 2020; Serafini, 2017). However, new advancements in statistics software and data analysis techniques such as GAMM are enriching the CDST toolbox. Other techniques that have been proposed as appropriate methods to study complex systems with a relation-intensive element are latent growth curve modeling (LGCM) and multilevel modeling (MLM) (Hiver & Al-Hoorie, 2020a; MacIntyre et al., 2017). To expand our CDST toolbox of relation-intensive approaches, we could also utilize network analysis, an underexplored methodology in SLA research.

## Network analysis

Network analysis has become a popular technique for studying complex systems in the field of psychology. Readers should be aware that there are many different types of network models; network analysis can be performed on cross-sectional data from a relation-intensive perspective (Epskamp & Fried, 2019; Hevey, 2018), and also on longitudinal time-series data from a time- or time-relation intensive perspective (Bringmann et al., 2013). Although we outline some other variants of network analysis later in the discussion section, it is beyond the scope of this article to describe each type of network analysis in detail. We have opted to focus on psychological networks with cross-sectional data for relation-intensive approaches, which is an underexplored dimension of CDST research in applied linguistics.

As readers may be more familiar with social network analysis, we would also like to briefly explain some differences between social networks and psychological networks. Social networks show patterns of relationships among individuals or groups, whereas psychological networks show patterns of relationships among variables (at item level or composite level). It is important to note that with social networks, the relationships between variables are *known*; social networks are created from an adjacency matrix, whereby the relationships between variables are directly observed (O'Malley & Onnela, 2019). In contrast, with psychological networks, relationships between variables are not known but are *estimated*. Psychological networks are estimated from a variance-covariance matrix, based on the strength of partial correlations between variables (Epskamp & Fried, 2019).

Psychological network analysis has been used to model constructs such as intelligence (van der Maas et al., 2006, 2017) cognitive development (Kievit, 2020), and mental disorders (Borsboom, 2017) from a CDST perspective, and has also been applied to clinical research on psychological disorders such as depression and eating disorders (Elliott et al., 2020; Lutz et al., 2018). In network models, variables (also referred to as components) are represented as circles called *nodes*. In psychological networks, nodes represent elements of a construct or an entire construct, such as attitudes or symptoms of a mental disorder. Lines between nodes are called *edges*, which represent the direct association between a pair of nodes. The strength of association between nodes is called the *edge weight*; the thicker the edge, the stronger the association. Edges in psychological networks are typically *undirected*, which reflect the hypothesized multicausal relationships between system components. Positive relationships are typically denoted using blue edges, while red edges are used to indicate negative relationships. The layout of the network model can be selected by the researcher. Psychological networks are often plotted (by default) using the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991), which places nodes with stronger connections closer together, and nodes with weaker connections further apart. Besides visual inspection, network models can be analyzed on several different levels, depending on what the research questions are. For example, researchers typically analyze the network density if the overall interest is the network structure or focus on particular nodes and edges (Burger et al., 2022).

The most common models used to estimate psychological networks are pairwise Markov random field (PMRF) models. Within PMRF models, Gaussian graphical models (GGM) are used with continuous multivariate data to estimate partial correlations between variables (Epskamp, 2014). Partial correlation networks are undirected graphs, estimated by analyzing the strength of correlations between variables after controlling for the effect of other measured variables in the network (Hevey, 2018). As



such, a psychological network can be viewed as a “nomological net, which functions as a specification of the phenomenological concepts or theoretical constructs of interest in a study, their observable manifestations, and the linkages between them” (Hiver & Al-Hoorie, 2016, p. 747). Psychological networks created using cross-sectional data can therefore be viewed as a snapshot of the system at a given time.

### Network analysis and CDST

Network analysis has some advantages over other relation-intensive methods used in CDST research. One advantage is that network analysis is more conceptually aligned with CDST compared to factor-based statistical techniques that are rooted in latent variable theory (Fried, 2020). Originally developed by Spearman (1904), factor models function under the theoretical assumption that a latent construct, such as intelligence or personality, can be measured through observable indicators (e.g., behavioral tests or questionnaire items). This means that there is a hypothesized unidirectional relationship from the latent construct to the observable indicator, whereby answers to questionnaire items or tests are thought to “reflect” the latent construct (Edward & Bagozzi, 2000). In contrast, from a network perspective, psychological constructs “exist as *systems* where components mutually influence each other without the need to call on latent variables” (Guyon *et al.*, 2017, p. 2). Statistically, factor models and psychological networks are closely related, as both analyze the covariance between observed variables. The difference between each approach is their competing causal explanations (Fried, 2020). As van Bork *et al.* (2019, p. 1) explain, “whereas latent variable approaches introduce unobserved common causes to explain the relations among observed variables, network approaches posit direct causal relations between observed variables.”

These two competing causal explanations are reflected in the choice of statistical model selected by the researcher. For example, factor-based techniques such as SEM or LGCM generate directed graphs, with edges from the latent construct to the observed indicators and/or between latent constructs, which are determined by the researcher *a priori*. Psychological network analysis is a more data-driven approach and produces an undirected graph with edges estimated between all nodes, better reflecting key CDST concepts such as multicausality and interconnectedness. This has already been noted in the field of psychology, where researchers working from a CDST perspective are using network analysis as an exploratory tool to better visualize the complex patterns of relations between variables of interest (Hilpert & Marchand, 2018; Sachisthal *et al.*, 2019; van der Maas *et al.*, 2017).

In this article, we explore how network analysis could be used to model psychological constructs that influence language learning from a relation-intensive perspective. We provide two examples of psychological networks created using the datasets of existing studies that are publicly available online in support of Open Science practices. As the nested nature of educational phenomena can be analyzed at multiple levels (Marchand & Hilpert, 2018), our network models illustrate two different levels of analysis; with nodes at item level and composite level. The first example is a network model of L2 motivation made using the dataset from Hiver and Al-Hoorie’s (2020b) study on the role of vision in L2 motivation. This example explores how an individual difference construct such as L2 motivation can be modeled as a complex system, by analyzing relationships between the L2MSS at the item level. The second example is a network model of individual differences in native language ultimate attainment, made using the dataset from Dąbrowska’s (2018) study. The second example takes a wider

relation-intensive perspective by analyzing interactions between multiple individual difference constructs at the composite level. Note that the authors of the original studies (Dąbrowska, 2018; Hiver & Al-Hoorie, 2020b) did not position their research within a CDST paradigm, and our reanalysis of their data is not a critique on their work.

We performed all statistical analyses using the open-source software R (R Core Team, 2020) and the R-packages *qgraph* (Epskamp et al., 2012) and *bootnet* (Epskamp et al., 2018a) in particular. The R code that we used to create these two examples is available in the online Supplementary Materials on our Open Science Framework (OSF) page. This article is not intended to serve as a tutorial in network analysis (for tutorials, we refer readers to Burger et al., 2022; Epskamp et al., 2018a; and Hevey, 2018). Rather, our overall aim is to raise awareness of this methodology and illustrate how it can be applied to model psychological constructs related to language learning from a relation-intensive CDST perspective. Within each example, we evaluate (a) the extent to which a network analysis of the datasets supports the same conclusions as the original authors and (b) whether network analysis can offer any additional insights to the original analyses.

### Example 1

The first example was made using the dataset from Hiver and Al-Hoorie's (2020b) study "Reexamining the Role of Cision in Second Language Motivation: A Preregistered Conceptual Replication of You, Dörnyei, and Csizér (2016)." Both Hiver and Al-Hoorie (2020b) and You et al. (2016) used SEM to explore interrelationships between components of the L2 Motivational Self System (L2MSS). The L2MSS is a theoretical paradigm that was developed by Dörnyei (2005, 2009) based on Possible Selves Theory (Markus & Nurius, 1986). The L2MSS is comprised of three components, each theorized to be a primary source of motivation to learn an L2: the Ideal L2 Self, the Ought-to L2 Self, and L2 Learning Experience. The ideal L2 self refers to learners' internal desires and wishes to learn the L2, while the ought-to L2 self refers to learner's perceived external duties and social pressures to learn the L2 (Dörnyei & Chan, 2013). L2 experience concerns learners' attitudes toward learning, based on their experience of the learning process and environment. In addition to these three components, vision and imagery are also considered key aspects of the L2MSS, whereby motivation is viewed as "a function of the language learners' vision of their desired future language selves" (Dörnyei & Chan, 2013, p. 437). Vision can be considered as a combination of imagery capacity and ideal selves and is typically measured by visual and auditory learning style preferences, and vividness of imagery capacity (You et al., 2016). A number of studies have used SEM to explore the interrelationships between these motivational constructs and the extent to which the L2MSS can predict language learning or intended effort (Dörnyei & Chan, 2013; Hiver & Al-Hoorie, 2020b; You et al., 2016). However, as You et al. (2016, p. 97) have pointed out, "because the L2 Motivational Self System was originally proposed as a framework with no directional links among the three components, past empirical studies employing SEM have not been uniform in specifying these interrelationships." For example, whereas some studies have presented a directed pathway from the ideal L2 self to L2 learning experience, other studies have reversed this relationship (for further details see You et al., 2016).

Hiver and Al-Hoorie (2020b) conducted a conceptual replication and extension of You et al. (2016) to evaluate the role of vision in L2 motivation and to assess whether intended effort is an outcome or a predictor of motivation. They justified these aims in

part due to the fact that You *et al.* did not test equivalent or competing models, which could be considered a form of confirmation bias. Hiver and Al-Hoorie also stressed the need for more robust research designs, and further replication of research on language motivation. Hiver and Al-Hoorie (2020b) collected data from 1297 L2 learners of English in secondary schools in South Korea. In addition to the same 10 scales of motivation and vision used by You *et al.*, Hiver and Al-Hoorie also included two measures of L2 proficiency, midterm grades and final exam grades, which were analyzed as one variable called L2 achievement. To determine the number of underlying factors, they submitted the dataset to Mokken scaling analysis, confirmatory factor analysis, exploratory factor analysis, scree plot, optimal coordinates, and parallel analysis (Hiver & Al-Hoorie, 2020b, p. 73). These analyses resulted in only four factors: visual style, ideal L2 self, ought-to L2 self, and intended effort. With these four factors and the measures of L2 achievement, Hiver and Al-Hoorie used SEM to test two competing causal models of vision and L2 motivation, where intended effort was either an antecedent or an outcome of motivation. Contrary to You *et al.* (2016), Hiver and Al-Hoorie hypothesized intended effort to be an antecedent of the ideal L2 self and the ought-to L2 self. In both competing models, vision (visual style) was considered a predictor of motivation, which was the same as You *et al.* (2016). Results showed that the model with intended effort as a predictor of motivation showed a better overall fit. Although this was contrary to You *et al.*'s model, Hiver and Al-Hoorie note that as their dataset and analyses differed greatly from the initial study, their model cannot be used to contradict You *et al.*'s model and call for further replication of research on the L2MSS.

In both studies (Hiver & Al-Hoorie, 2020b; You *et al.*, 2016), the authors were interested in the relationships between the L2MSS, vision, and intended effort. By using SEM, they operationalized motivational constructs as latent variables, depicting hypothesized causal relationships between latent constructs with unidirectional arrows. However, in both studies, the authors note potential issues and limitations of using SEM to model interactions between motivational constructs. One issue relates to the theorized dynamic nature of the L2MSS and the multicausal relationships between motivational constructs. Possible Selves Theory was originally proposed to have dynamic qualities, whereby current and ideal selves are shaped by multiple ongoing processes (Henry, 2014; Markus & Nurius, 1986). For example, Hiver and Al-Hoorie speculate that once an L2 learner puts in the effort and engages in the L2 learning process, "there will be a dynamic interaction between motivation ... and task demands, leading to continuous recalibration of that motivational construct" (2020b, p. 86). One might question the extent to which SEM can effectively model these dynamic interactions, as SEM operationalizes motivational constructs as latent variables with a unidirectional causal relationship. In fact, both studies' authors acknowledge that a further limitation of SEM is that it requires the researcher to specify the direction of the relationship between latent constructs. SEM can only test the theoretical model that is selected by the researcher, although equivalent or alternative models may likely exist. This issue was illustrated by Hiver and Al-Hoorie's (2020b) two competing SEM models. As discussed earlier, there has already been discussion of how the L2MSS could be conceptualized as a complex system (Henry, 2014, 2017) and manifest as an attractor state (Henry, 2017; Hiver, 2014; Waninge *et al.*, 2014). From a CDST perspective, causal relationships between motivational constructs are not unidirectional, but reciprocal. To further investigate the relationship between motivation and intended effort, Hiver and Al-Hoorie (2020b) have encouraged researchers to consider using nonrecursive models where causality is reciprocal. In this first example, we illustrate how network analysis

can be used to model the L2MSS as a complex system, with hypothesized reciprocal causation between motivational constructs with nodes at item level.

### Network estimation and visualization

Figure 1 is a GGM of the L2MSS that we made with the dataset from Hiver and Al-Hoorie’s (2020b) study. In support of open science practices, they made their dataset and analyses publicly available through the OSF website. To allow for ease of comparison, we included the same variables in our network analysis as Hiver and Al-Hoorie’s SEM analyses, with the exception of visual style 2, which we explain in a later section. The network model in Figure 1 has nodes at item level, to better explore the interrelatedness of these motivational constructs, and the questionnaire items used to measure them. Table 1 contains information about which items correspond to each node.

We chose GGM model selection (ggmModSelect function implemented in the *bootnet* R-package; Epskamp et al., 2018a) as estimation method because of the large size of the dataset. Model search works by setting edges to zero and using a stepwise algorithm to continuously estimate the model until the optimal model is identified (Epskamp, 2014). This technique uses Bayesian information criterion (BIC) obtained through estimating the maximum likelihood of sparsity.

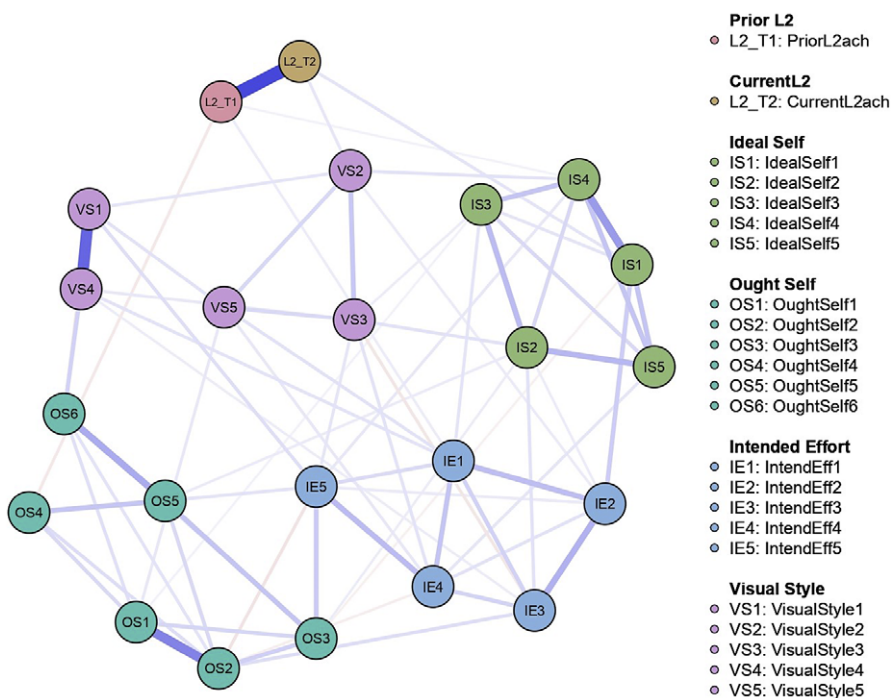


Figure 1. A network model of the L2MSS and L2 achievement.

Note: In this network model of the L2MSS, there are four motivational constructs: the ideal L2 self, the ought-to L2 self, intended effort, and visual style. Each node represents a questionnaire item. Ought-to L2 self has been measured with six questionnaire items, and the other motivational constructs with five questionnaire items. There are also two composite measures of L2 proficiency: L2\_T1 (students’ mid-term grades) and L2\_T2 (students’ final grades).

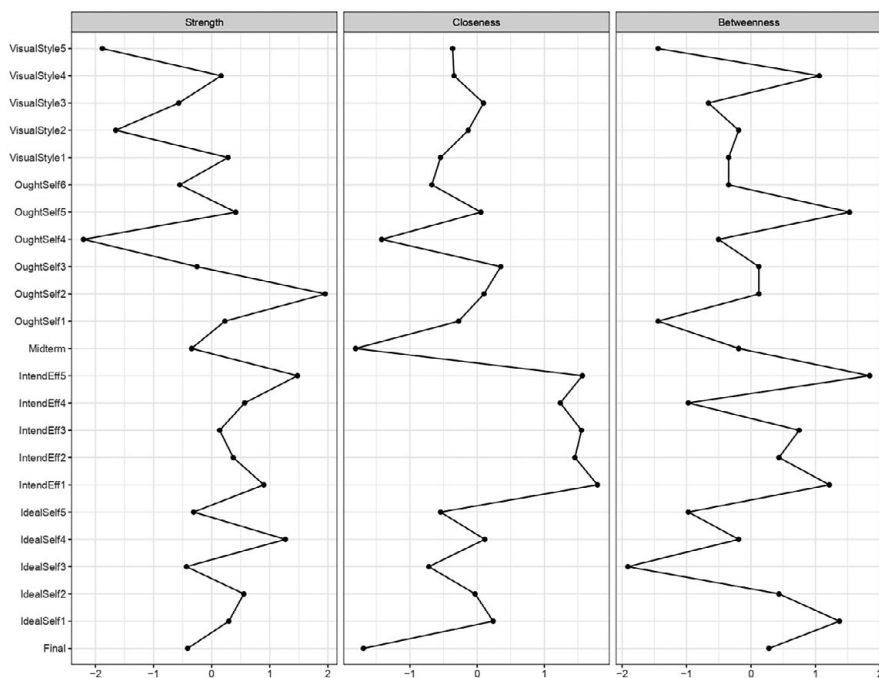
**Table 1.** Legend of node labels

Node labels	Items
L2 achievement	
L2_T1	Mid-term grades
L2_T2	End of term grades
Ideal L2 self	
IS1	I can imagine myself speaking English in the future with foreign friends at parties
IS2	I can imagine myself in the future giving an English speech successfully to the public
IS3	I can imagine a situation in which I am doing business in foreigners by speaking English
IS4	I can imagine myself speaking English in the future having a discussion with foreign friends in English
IS5	I can imagine that in the future in a café with light music, a foreign friend and I will be chatting in English casually over a cup of coffee
Ought-to L2 self	
OS1	Studying English is important to me to gain the approval of my teachers
OS2	Studying English is important to me to gain the approval of my peers
OS3	Studying English is important to me to gain the approval of the society
OS4	I study English because close friends of mine think it is important
OS5	I consider learning English important because the people I respect think that I should do it
OS6	My parents/family believe that I must study English to be an educated person
Visual style	
VS1	I use color coding (e.g., highlighter pen) to help me as I learn
VS2	Charts, diagrams, and maps help me understand what someone says
VS3	When I listen to a teacher, I imagine pictures, numbers, or words
VS4	I highlight the text in different colors when I study English
VS5	I learn better by reading what the teacher writes on the board
Intended effort	
IE1	I am prepared to expend a lot of effort in learning English
IE2	I find learning English really interesting
IE3	I would like to concentrate on studying English more than any other topic
IE4	Even if I failed in English learning, I would still learn English very hard
IE5	English would still be important to me in the future even if I failed in my English course

In their original analyses, Hiver and Al-Hoorie (2020b) tested for normality and found that their data were not multivariate normal both in skewness and kurtosis. For this reason, the network was estimated using Spearman correlations (Epskamp, 2014).

After estimating the model, we evaluated the stability of the network structure in terms of edge-weight accuracy using bootstrapping (see Epskamp et al., 2018a for an in-depth explanation of bootstrapping in psychological networks). We used 5,000 samples of the nonparametric bootstrap to assess the variability of the edge-weights. This step should always be performed (Epskamp et al., 2018a) as any interpretation of the network becomes limited if the network is unstable (Burger et al., 2022). The results show a good overlap between the estimated model and the bootstrapped edge-weights, indicating that the network of Figure 1 is stable. The results of the nonparametric bootstrap can be viewed in the supplementary materials.

To assess the stability of the centrality coefficients, we again used bootstrapping. We used the case-dropping bootstrap, specifically developed to this aim (Epskamp et al., 2018a). The case-dropping bootstrap assesses the stability of the order of centrality in subsets of the data, that is, after systematically dropping an increasing percentage of participants from the dataset. The centrality stability for “strength” centrality was



**Figure 2.** Centrality plots for the L2MSS network.  
 Note: Centrality plots for the network model of the L2MSS. Centrality measures are shown as standardized z-scores. The raw centrality indices can be found in the online Supplementary Materials.

estimated on a sample of 5,000 bootstraps, which resulted in a correlation stability coefficient (CS-coefficient) of 0.52 for the “strength” centrality. This is above the 0.5 (CS-coefficient) recommendation (Epskamp et al., 2018a), which is why we conclude that the stability of node centrality in this network model is good. The results are presented in the supplementary materials.

Based on the centrality indices (see Figure 2), ought-to L2 self 2 is the most central component in the network model in Figure 1 in terms of node strength, followed by intended effort 5. The questionnaire items that correspond to these components are “Studying English is important to me to gain the approval of my peers” and “English would be still important to me in the future even if I failed in my English course.” This suggests that peer approval and perceived future importance of English play important roles in L2 motivation, as they are the strongest direct relationships with other motivational constructs in the system.

In addition to node strength, we also computed node centrality indices based on closeness and betweenness. The closeness index “indicates a short average distance of a specific node to all other nodes” (Hevey, 2018, p. 311). In the network model, the nodes with the highest closeness are the five intended effort nodes. This is an interesting find and indicates that intended effort may have an integral role in L2 motivation. Although the role of central components is not yet fully understood, it is thought that central nodes with high closeness are the most likely to both effect changes and be affected by changes in the system (Hevey, 2018). The third measure of centrality, betweenness, refers to how well one node connects other nodes together; nodes with high

betweenness lie on the shortest path between pairs of nodes. As shown in [Figure 2](#), the node with the highest betweenness is intended effort 5, followed by ought-to L2 self 5 and ideal L2 self 1.

### *Interpreting the network model*

Both You *et al.* (2016) and Hiver and Al-Hoorie (2020b) used SEM to evaluate the relationships between the L2MSS, vision, and intended effort. The network model contains four motivational constructs (ideal L2 self, ought-to L2 self, intended effort, visual style) and the two measures of L2 achievement (midterm grades and final exam grades). We can see the wider interconnectedness of components in the system, with multiple interactions across different motivational constructs.

One of the first things we notice when looking at this network model is that, although the motivational constructs are interrelated, there are only a few weak edges between any of the motivational constructs and the two measures of L2 proficiency. For example, final grades have a weak partial correlation with ideal self 1 (0.10) and midterm grades have a weak negative partial correlation with ought-to self 4 (−0.09). The network analysis results are consistent with Hiver and Al-Hoorie's (2020b) study, where the ideal L2 self was only a weak predictor of L2 achievement (accounting for less than 1% of the variance), and the ought-to L2 self had almost no predictive value.

### *Visual style*

The visual style scale consists of five questionnaire items. As can be seen in [Figure 1](#), although each of the five measures of visual style are grouped together, the nodes are not as closely grouped together compared to nodes measuring other constructs. In Hiver and Al-Hoorie's SEM analyses, they excluded visual style 2 to improve convergent validity, and also note that this scale had the lowest reliability in You *et al.*'s (2016) study. Removing items is typical with latent variable approaches, where researchers drop variables that do not load onto factors or if there are cross-loadings (Fried, 2020). With network analysis however, Fried (2020, p. 21) has pointed out that "items that load onto two factors simultaneously make for the potentially most interesting items because they may build causal bridges between two communities of items." Because of this, we decided to include visual style 2 in the network analysis. The network model shows that visual style 2 is linked to three other nodes measuring visual style, and also has a weak partial correlation with one measure of L2 achievement, one measure of the ideal L2 self, and one measure of intended effort. While visual style 2 was left out of the SEM analyses, results of the network analysis tentatively suggest that this questionnaire item may function as a bridge node between other motivational constructs. In both You *et al.* (2016) and Hiver and Al-Hoorie's (2020b) SEM analyses, they treated visual style as a predictor of the ideal L2 self and the ought-to L2 self. The network model is an undirected graph, so our analyses cannot provide additional insights into whether visual style is a predictor or outcome of motivation. What the network analysis does provide, is a more complex pattern of relationships between visual style and other system components than the original analyses. The nodes that measure visual style are partially correlated with components from all other motivational constructs in the network, as well as one measure of language achievement.

### Intended effort

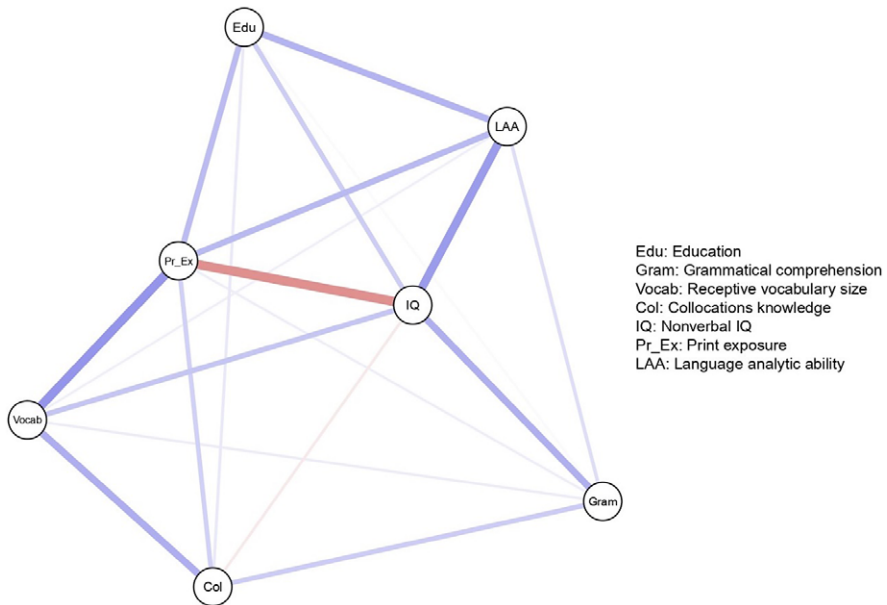
Besides the role of vision, You et al. and Hiver and Al-Hoorie were also interested in the direction of the relationship between intended effort and the L2MSS. Hiver and Al-Hoorie's analyses of two competing SEM models showed that intended effort was a better predictor of the ideal L2 self and ought-to L2 self than an outcome. Previous research has provided empirical evidence for reciprocal causal relationships between motivation and academic achievement (Vu et al., 2021). The network model shows that components of intended effort are related to components of all other subsystems, as well as L2 achievement, indicating a complex pattern of relationships. The results of the centrality indices highlight the overall importance of intended effort in L2 motivation, as the five intended effort variables have the highest closeness index in the network. Overall, intended effort 5 emerges as the most central component of the network. This item refers to the statement "*English would be still important to me in the future even if I failed in my English course.*" Intended effort 5 also has the highest centrality in terms of betweenness, and the second highest in terms of closeness and strength. The question surrounding the role of central components will be further discussed later in this article.

### Example 2

The second example illustrates how network analysis can be used to explore the relationships between multiple individual differences using the dataset from Dąbrowska's (2018) study *Experience, aptitude and individual differences in native language ultimate attainment*. The dataset is publicly available online using the IRIS Database. The network model made from Dąbrowska's (2018) dataset presents a different level of analysis from the previous example. In contrast to the network model in example 1, where each node represents a single questionnaire item, each node in the network model in Figure 3 represents a distinct variable measured by aggregated task scores. In the original study, Dąbrowska tested the assumption that adult native speakers tend to converge on the same grammar. She addressed this question by considering two opposing approaches to language acquisition: the usage-based perspective and the modular perspective. From a usage-based perspective, language abilities are thought to emerge out of interactions between general cognitive mechanisms and exposure to linguistic input (Ellis & Wulff, 2018). From this perspective, "causal mechanisms interact iteratively to produce what appears to be structure" (Bybee & Beckner, 2009, p. 23). A usage-based approach is thus aligned with CDST, where linguistic knowledge emerges as a network of interrelated and interacting components. In contrast, from a modular perspective, language abilities are thought to stem from an innate universal grammar, whereby different types of language knowledge rely on autonomous modules within the mind (Tan & Shojamanesh, 2019).

Dąbrowska (2018) discusses the plausibility of these two theories in connection with analyses of a dataset of 90 native English speakers' performance on different linguistic and nonlinguistic tasks. She first analyzed the amount of individual variation on six tasks that measured grammatical comprehension, receptive vocabulary, collocations, nonverbal IQ, language analytic ability, and print exposure. Full details regarding which tests were used to measure each construct can be found in the original study. Dąbrowska then conducted Pearson correlations to explore interactions between the six aforementioned tasks as well as education (measured by number of years spent in education). This revealed several significant correlations between the measures of





**Figure 3.** A network model of individual differences in native language ultimate attainment. *Note:* The nodes in this network are composite scores representing three measures of language proficiency and four individual differences. The three proficiency measures are receptive vocabulary, collocations, and grammatical comprehension. The four individual differences are nonverbal IQ, print exposure, language analytic ability, and years of education.

language knowledge as well as between other variables. To determine potential causes of individual differences in linguistic knowledge, Dąbrowska then conducted regression analyses with the four predictor variables (nonverbal IQ, language analytic ability, print exposure, and education) on each of the three measures of language knowledge. Overall, results showed that nonverbal IQ was strongly related to grammar and vocabulary, but not to collocations. Language analytic ability was also significantly related to grammar and vocabulary, as well as several other variables. Print exposure contributed more to vocabulary and collocations than to grammar, and education only weakly predicted each measure of language knowledge. Based on the significant correlations between the three measures of language knowledge and the fact that the same nonlinguistic variables predicted different areas of language knowledge, Dąbrowska concluded that these findings support a usage-based approach.

### **Network estimation and visualization**

Figure 3 is a GGM of partial correlations that includes the same seven variables used in Dąbrowska's analyses. The network model was estimated using the "least absolute shrinkage and selection operator" (LASSO), which is considered an appropriate estimation method for smaller datasets (Epskamp *et al.*, 2018a; Hevey, 2018). The LASSO technique results in a sparser network, using only a relatively small number of edges to explain the covariance in structure (Epskamp *et al.*, 2018b). This makes the estimated model more interpretable and accurate, as very small edges are removed from

the estimated network (Epskamp et al., 2018a; Hevey, 2018). The LASSO applies a regularization technique that is controlled by a tuning parameter. The tuning parameter was selected by minimizing the Extended Bayesian Information Criterion (EBIC), for which we used the default setting of 0.5.

To assess network stability, we used a nonparametric bootstrap of 5,000 samples. Bootstrapping results can be found in the supplementary materials. The bootstraps show wide 95% confidence intervals, meaning that the estimated network structure is not very stable and the found links should be interpreted with care. As such, our discussion and interpretation of this network model is tentative, and a larger sample size is needed to draw any strong conclusions. We did not compute centrality indices for this dataset because the aim of this network analysis was to explore overall patterns of relationships between variables, and also given the small number of variables in this model.

### *Interpreting the network model*

The network model in Figure 3 illustrates a complex system of interdependent relationships between linguistic and nonlinguistic variables. Each node in the network model in Figure 3 represents a composite variable. For example, the node “collocations” consists of 40 multiple choice items on the Words That Go Together test, and the node “print exposure” consists of 130 items on the Author Recognition Test. From a CDST perspective, the network model in Figure 3 provides a visualization of how different aspects of language knowledge are related to both internal resources (nonverbal IQ and language analytic ability) and external resources (print exposure and education). When comparing to the results of the regression analyses in the original study, the network model reflects the same overall patterns of relationships between individual differences in language knowledge. For example, nonverbal IQ is more strongly associated with grammar and vocabulary than with collocations, and print exposure is more strongly associated with vocabulary and collocations than with grammar. The fact that both analyses reveal the same overall patterns is not surprising because partial correlations and multiple regression coefficients both estimate of the strength of relationships between variables while controlling for the effects of other measured variables (Hevey, 2018). The key difference is that regression analysis imposes unidirectional causal relationships between specific variables selected by the researcher, whereas with network analysis there are no assumptions regarding the direction of the relationships.

There are some subtle differences between the results of the network analyses and Dąbrowska's analyses. For instance, whereas Dąbrowska found that language analytic ability was significantly related to both grammar and vocabulary, the network analysis shows that language analytic ability is only very weakly associated with vocabulary. In the network model, the relationship between language analytic ability and vocabulary knowledge appears to be altered by print exposure and nonverbal IQ. Similarly, while Dąbrowska's analyses showed that education weakly predicted each measure of language knowledge, the network analysis shows that the relationship between education and language knowledge becomes weaker after controlling for the effects of print exposure, nonverbal IQ, and language analytic ability. It is also interesting to note that in the network model, the negative relationship between print exposure and nonverbal IQ becomes stronger after controlling for other variables. Another minor difference is that network analysis revealed a negative relationship

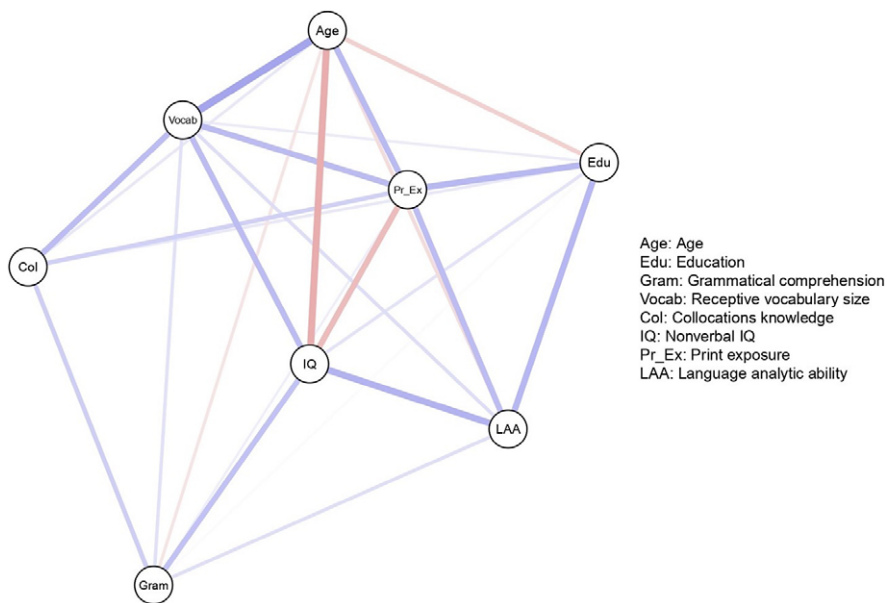
between nonverbal IQ and print exposure whereas in Dąbrowska's analyses this relationship was positive. The reason for this difference is that Dąbrowska transformed the raw IQ scores into percentages, while we opted to conduct analyses with the raw IQ scores. To confirm this, we conducted Pearson correlations between print exposure and both the raw and transformed IQ scores that showed that print exposure had a weak negative correlation with raw IQ scores ( $r(88) = -.03$ ,  $p = .719$ ) and a weak positive correlation with transformed IQ scores ( $r(88) = .08$ ,  $p = .440$ ). However as these are very small differences, they cannot be interpreted as meaningful. These slight differences revealed by the network analysis could be due to the fact that we included all seven variables in the network analysis, whereas Dąbrowska conducted three separate regression analyses for each measure of language knowledge. By taking a more holistic approach including all variables within the same analysis, additional patterns of relationships were revealed. This then raises the question of how many variables should be included when working from a CDST perspective.

### *Adding age to the network model*

To explore this idea, we expanded on the original study by adding the variable "age" to the network model. Participants' ages were contained within the original dataset that is available online, but Dąbrowska did not include this variable in her analyses. It seemed particularly interesting to include this variable because Dąbrowska used the dataset to evaluate the usage-based approach and the modular approach to language acquisition. Age is an indirect measure of language experience. With first language development, it is logical to assume that the older a person is, the more exposure to linguistic input they have. Thus, from a usage-based perspective, we might hypothesize age to be significantly related to a number of other variables, including measures of language knowledge. The 90 participants in Dąbrowska's study varied greatly in age, with a range of 17 to 65 and a mean age of 38. The network model in Figure 4 is a GGM of partial correlations between eight variables (the seven variables from the original analyses plus age). The model was made following the same procedures described for the network model in Figure 3. Similarly to the model without age, the bootstraps show wide 95% confidence intervals, meaning that the estimated network structure is also not very stable and the found links should be interpreted with care.

The network model in Figure 4 shows that age is partially correlated with all other variables. Out of the three measures of language knowledge (grammar, vocabulary, and collocations), age is most strongly linked to vocabulary (0.35), which is the strongest positive edge in the network.<sup>1</sup> This is in line with previous studies which have shown that vocabulary is typically the only aspect of language knowledge that does not tend to decline with age (Reifegerste, 2021). As could be expected, age is also related to print exposure. Age has a negative association with nonverbal IQ and language analytic ability, which is consistent with previous research on cognitive decline and aging (Reifegerste, 2021). There is also a negative relationship between age and education,

<sup>1</sup>We conducted the *bootstrapped difference test* to check whether the edges in the network significantly differ from each other, in the supplementary materials. The edge Age-Vocabulary is significantly stronger than the edge Age-Grammar, but not from the edge Age-Collocations knowledge. This means that the difference between these edges has to be interpreted with care.



**Figure 4.** A network model of individual differences in language knowledge, including age.  
 Note: In addition to the same variables as the network model in Figure 3, this model also has the variable age. Blue edges denote positive partial correlations and red edges denote negative partial correlations.

which is logical considering that the percentage of university attendance has increased over the years. Overall, the considerable effect that age has on the system provides tentative support for the usage-based approach. Using network analysis, we can visualize individual differences in language abilities as a complex system. While we cannot draw conclusions about emergent processes from cross-sectional data, based on this network model, we could speculate that vocabulary knowledge emerges out of interactions between cognitive abilities (nonverbal IQ) and other language experience (print exposure) throughout the lifespan (age).

In addition, controlling for age alters the partial correlations between other nodes. For example, in the model that includes age, vocabulary knowledge has a weak positive relationship with grammar (0.11) and language analytic ability (0.12), whereas in the model without age, these relationships are weaker (0.06 and 0.05). This indicates that age is a moderating variable. When controlling for age, the edge weight between grammar and language analytic ability is stronger (more positive) because age has negative partial correlations with grammar and language analytic ability. In a similar way, age also moderates the relationship between IQ and print exposure; these variables have an edge weight of  $-0.43$  without age, and  $-0.26$  when age is added to the model. In this case, the edge weight between IQ and print exposure is weaker (less positive) when controlling for age because age has negative partial correlations with IQ and print exposure.

Although the network models estimated in example 2 are not stable, and a larger sample size is necessary to draw any firm conclusions, our examples serve to illustrate how network analysis can be used to model multiple individual differences in language learning from a CDST perspective. The network analyses support the same conclusions

as the original study (Dąbrowska, 2018), but rather than analyzing unidirectional relationships between individual difference constructs, the undirected network models in example two depict hypothesized multicausal relationships between variables. By estimating partial correlations between all variables, network analyses also reveal a more complex network of relationships between variables than the original study's regression analyses, offering additional insights into the data.

## Discussion

We have provided two examples of how network analysis can be used to model complex systems from a relation-intensive perspective. These examples serve to illustrate how a network approach can offer new insights into which components form a system and the nature of the relationships between components. Network analysis is conceptually aligned with CDST, enabling us to model hypothesized multicausal relationships between variables. We have shown how network analysis of cross-sectional data can be used to model individual difference constructs as complex systems, viewing the network as a snapshot of (part of) a system in time. We illustrated this in example 1 with nodes at item level, to analyze motivational constructs on a micro level, and in example 2 with node at composite level, to analyze the relationships between individual differences and language knowledge on a more macro level. In both examples, network analysis complements the original analyses by providing a more intricate pattern of relationships between system components, and deeper understanding into the variables of interest.

Besides the examples of psychological network analysis in this article, there are other applications of network analysis that could also be beneficial to SLA researchers, such as the network comparison test and dynamic network analysis. It is also important to acknowledge that psychological network analysis is still a relatively new statistical technique, and there are some unanswered questions regarding how certain aspects of CDST fit with network analysis, for example regarding the question of how many variables to include and the role of central components. In the following section, we discuss some of these questions and highlight additional applications of network analysis that could be applied to SLA research.

### *The network comparison test*

The network comparison test is an application of network analysis that can be used to compare group differences. The network comparison test statistically compares the networks of two (or more) groups, such as in terms of node centrality and global strength (van Borkulo *et al.*, 2022). Networks can also be compared visually, which is typically done by constraining the layout of the two models for ease of visual comparison. Blanco *et al.* (2020) used a network comparison test to compare the effects of two different interventions on treating depression. One group of patients ( $n = 45$ ) received a 10-week Positive Psychology Intervention (PPI) while another group ( $n = 48$ ) received a 10-week Cognitive-Behavioral Therapy (CBT) program. Both groups completed clinical assessments of depression symptoms before and after the intervention treatments. Blanco *et al.* (2020) used this data to create two network models to compare before and after treatment. Results of the network comparison test showed that only the PPI group showed significant changes in several edge weights and global strength after intervention. In SLA research, the network

comparison test could be used to statistically compare the networks of learners at different proficiency levels, at different time points, or across learning conditions. Both You et al. (2016) and Hiver and Al-Hoorie (2020b) conducted additional analyses to compare the roles of vision and intended effort across male and female L2 learners. This comparison could be also done using a network comparison test. As such, the network comparison test could strengthen CDST inspired research by providing a means for hypothesis testing and generalizations. Comparing networks across groups could also help to ascertain the phenomenological validity of conceptualizing abstract psychological phenomena as complex systems. In addition, the network comparison test could provide insight into how to influence systems' behavior, as illustrated by Blanco et al. (2020), and could a useful tool for SLA researchers considering complex interventions (Hiver et al., 2022).

### *Dynamic network analysis*

In examples 1 and 2, we took a relation-intensive CDST approach by estimating GGMs of cross-sectional data. The GGM can also be used with time-intensive and time-relation intensive research designs, for single subjects and group data, respectively. Dynamic network analysis requires intensive repeated measurements of variables, such as with a time-series or panel design, typically obtained through Experience Sampling Method (ESM), whereby participants provide self-reports at regular intervals during the day (Bringmann et al., 2013). With single-subject data, auto regressive (AR) modeling can model time dynamics within an individual by regressing one variable on a previous measurement of the same variable (called a *lagged* variable). The vector auto regressive (VAR) model is the multivariate extension of the AR model, where "a variable is regressed on all the lagged variables in the dynamic system" (van Bork et al., 2018, p. 18). The VAR model has two extensions: graphical VAR and multilevel VAR. For single-subject data, graphical VAR can be used to create both temporal and contemporaneous networks using the GGM (Epskamp et al., 2018). Temporal networks have directed edges and show how the state of variables at one time point influence the state of variables at the next time point. A contemporaneous network model shows how variables predict each other at the same measurement occasion, after accounting for temporal effects (Epskamp et al., 2018b), similarly to GAMMs and LCGMs. Multilevel VAR modeling can be used to model both within-group and between-group variance over time (Bringmann et al., 2013). For example, Bringmann et al. (2013) combined VAR and multilevel VAR to follow 129 participants' changes in depressive symptoms during a treatment intervention, modeling time dynamics at the individual and group level.

In the field of clinical psychology, researchers are exploring how dynamic network modeling could provide insight into how people develop disorders over time, with the aim of using this knowledge to target group and/or individual treatment interventions (Bringmann et al., 2013; David et al., 2018; van Bork et al., 2018). Dynamic network analysis could also prove to be a useful methodology for CDST researchers in applied linguistics, and a few SLA studies have used ESM. For example, Waninge et al. (2014) micro-mapped the motivational dynamics of four learners during their language lessons. They took measurements at 5-minute intervals throughout lessons, resulting in 10 observations per class. Similarly, Khajavy et al. (2021) used ESM to examine the dynamic relationships between willingness to communicate (WTC), anxiety, and enjoyment of 38 students throughout six language lessons. Students indicated their

level of WTC, anxiety, and enjoyment on a scale of 1 to 10 at 5-minute intervals, resulting in 10 observations per class. Gregersen *et al.* (2020) also used ESM to explore the dynamics of language teacher well-being, where teachers used an app to respond to a short survey 10 times a day for 7 days. Although smartphone technology has the potential to use ESM more easily than in the past (Arndt *et al.*, 2021; Gregersen *et al.*, 2020), it is still extremely challenging in most applied linguistics research settings to obtain a large enough number of observations to conduct a dynamic network analysis. For example, in the study by Bringmann *et al.* (2013), participants recorded depressive symptoms 10 times a day for 12 days, which resulted in a total of 120 observations per participant.

### **Nonlinearity**

GGMs and VAR models are estimated based on assumptions of multivariate normality that assume linear relationships between variables (Epskamp *et al.*, 2018b). As such, these models may not present a fully accurate view of the data if the relationships between variables are nonlinear. For cross-sectional data, the Ising model is a nonlinear model used for binary variables (Finnemann *et al.*, 2021), but nonlinear models for continuous variables have not yet been developed. For longitudinal data, while VAR models fit linear effects, new types of network analysis have been developed that can also capture nonlinear relationships between variables (Haslbeck *et al.*, 2021). The findings from several CDST studies with a time element have shown that language development, and its relationship with individual differences, is nonlinear (Fogal, 2022; Pfenninger, 2020; Piniel & Csisér, 2014). This is why some CDST research designs that include a time element are using techniques such as GAMM instead of LGCM, as GAMM can handle nonlinearity (Pfenninger, 2020). Researchers in the field of network psychometrics have recently combined the VAR model with a Generalized Additive Model (GAM) framework, to estimate time-varying VAR models (Haslbeck *et al.*, 2021). The field of network psychometrics is developing rapidly and is likely to produce other useful techniques in the future that could further enrich our methodological toolbox.

### **Latent network analysis**

In the first example of the L2MSS, we compared Hiver and Al-Hoorie's (2020b) SEM with our network analysis. For the past 100 years, psychological constructs have been studied using latent variable approaches, which assume that observed variables correlate because they reflect the same underlying construct (van Bork *et al.*, 2019). Network analysis has been put forward as an alternative to latent variable approaches. From a network perspective, correlations between observed variables may reflect mutual interaction between psychological processes (van der Maas *et al.*, 2006). Although these two approaches have different competing causal explanations for the covariance between observed variables, both create models for variance-covariance matrices and are thus statistically equivalent (van Bork *et al.*, 2019; van der Maas *et al.*, 2006). Because of this statistical equivalence, researchers have explored the idea that combining these two approaches could be complementary, resulting in *latent network analysis* (Golino & Epskamp, 2017; Guyon *et al.*, 2017). Conceptually, a combined approach assumes that manifestations of psychological attributes have a common cause (latent variables) and that these latent variables interact (as a complex system) (Guyon *et al.*, 2017). For

example, Epskamp et al. (2017) have used latent network modeling to explore the structure of interdependent relationships between latent variables. An advantage of latent network analysis is that due to the incorporation of factor-based statistical techniques, it is possible to test model fit against data, which is a limitation of psychological network analysis (Epskamp et al., 2017; van der Maas et al., 2017). It can also be considered a useful way of exploring latent variables within a dataset because clusters in the network can tell us about the factor structures present, without having to impose the direction of the relationship like SEM (Golino & Epskamp, 2017). Latent network analysis can be used with cross-sectional data as well as time-series and panel data.

### *The role of central components*

In example 1, we computed centrality indices for the network model of the L2MSS, which showed that the intended effort nodes have the highest centrality. Researchers from different fields have questioned whether central components have predictive ability and can be used to target interventions. The role of central components has so far provided insights into the dynamic processes of genetic networks, cortical networks, and ecosystems (for a detailed description see Rodrigues, 2019). In the field of clinical psychology, findings from few studies indicate that central components could be used to target treatment interventions and make predictions about diagnoses. For example, in clinical research on eating disorders, central components have been predictive of treatment dropout (Lutz et al., 2018) and treatment outcomes (Elliott et al., 2020). The idea behind using central nodes to target interventions is that these nodes are more likely to have bigger effects (either directly or indirectly) on the rest of the system compared to targeting a less central node (Rouquette et al., 2018). Nodes with high closeness in particular are more likely to be affected by changes in other components of the system and are also more likely to trigger change.

From the first network model example in Figure 1, intended effort had the highest node centrality in terms of closeness, suggesting that intended effort plays a key role in triggering the dynamic processes involved in L2 motivation. This fits with Hiver and Al-Hoorie's (2020b, p. 86) idea that putting in the effort to learn a language results in dynamic interaction between motivational constructs and task demands.

However, readers should note that the use of centrality indices in psychological networks is much debated (Bringmann et al., 2019). Centrality indices stem from social network analysis, whereby the relationship between components/nodes is known; the connections between nodes are observable. In comparison, in psychological networks the relationship between nodes is not directly observed, but is estimated, based on the strength of partial correlations between our measurements of psychological constructs. Bringmann and colleagues (Bringmann et al., 2019) have advised researchers to interpret centrality measures with care, especially betweenness and closeness centrality, as they are difficult to interpret and are often unstable.

### *The number of variables to include*

Complex systems are characterized by dynamic interaction between multiple internal and external subsystems (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008). However, given the theoretical and practical impossibilities of analyzing the complete



interconnected of a whole system, CDST researchers have to find a balance between oversimplification and undersimplification. Larsen-Freeman *et al.* (2011) have pointed out that a main methodological concern for CDST researchers is drawing boundaries and defining what we conceptualize to be a “functional whole.” Yet, when conducting network analysis, Hevey (2018, p. 307) has reasoned that it is “critically important to measure such potential confounding variables to ensure that their effects are controlled for.” The network model of Dąbrowska’s (2018) dataset in example 2 that includes age illustrates Hevey’s reasoning, as age moderates the relationships between other system components. It is highly likely that there are also other confounding variables that have been omitted from the model, such as socioeconomic status, L2 knowledge and experience, gender, and other cognitive abilities. As with other types of modeling, adding further variables to the network model could have both predictable and unpredictable effects on the rest of the system. Yet from a CDST perspective, it is theoretically impossible to measure every component of a system. What network analysis *can* do, is capture at least part of a system. Thus, while we acknowledge the potential of a network approach to SLA and individual differences, it is important to be mindful of its limitations.

### *Generalizability*

Generalizability is another debated topic in CDST research. Several researchers have pointed out the lack of generalizability of CDST studies and the lack of practical implications that CDST can currently offer to the field of applied linguistics (Hiver *et al.*, 2022; Palloti, 2022). Generalizability is a complex topic and is related to the distinction between idiographic and nomothetic methodological approaches. Idiographic approaches focus on the individual level with within-subject designs, analyzing intraindividual differences (Hamaker, 2012). Idiographic approaches use longitudinal data and process-focused analyses. In contrast, nomothetic approaches focus on the group level with between-subject designs, analyzing interindividual differences (Hamaker, 2012). Nomothetic approaches use cross-sectional data and product-focused analyses.

The majority of CDST studies to date have used idiographic approaches (Hiver *et al.*, 2022) because it is difficult to generalize from cross-sectional models to individual dynamics. This concept is known as the ergodicity problem: The idea that group statistics cannot be generalized to the individual and vice-versa (Lowie & Verspoor, 2019). As Molenaar (2004, p. 225) has pointed out, “only under very strict conditions—which are hardly obtained in real psychological processes—can a generalization be made from a structure of interindividual variation to the analogous structure of intraindividual variation.” However, this does not mean that idiographic and nomothetic approaches are in competition (Salvatore & Valsiner, 2010). In fact, they can be viewed as complementary, or two sides of the same coin (Grice, 2004). When discussing the idiographic-nomothetic debate in relation to research on personality, Grice (2004) argued that:

Establishing the uniqueness of some person’s developmental history, attitudes, thoughts, behaviors etc., would require the negation of nomothetic principles. Conversely, establishing the validity of a nomothetic principle that holds for all people would require the study of individual persons, not simply aggregates of

persons. A true study of personality is therefore necessarily idiographic *and* nomothetic. (p. 205)

Lowie and Verspoor (2019) have illustrated this point in relation to SLA, by investigating the role of motivation and aptitude in both a group study and in 22 longitudinal case studies. Their analyses showed that while learners showed different intraindividual learning trajectories over time, there were overall similarities between learners in terms of motivation and aptitude.

While some have argued that the idiographic approach undermines generalization (Palloti, 2022; Spencer & Schöner, 2003), others have argued that idiography is a way to pursue generalized knowledge (Salvatore & Valsiner, 2010). As Salvatore and Valsiner (2010) have claimed, idiography is “the pursuit of nomothetic knowledge *through the singularity of the psychological and social phenomena*” [emphasis in original] (p. 820). It is also important to note that nomothetic refers to what can be generalized across a sample population (e.g., from aggregated cross-sectional data), not what can be taken as a general law across all populations (Hamaker, 2012). Hence, as with any other cross-sectional data analysis, results of network analysis can only tell us about the population from which the data was sampled and cannot be taken as a general law across all populations or all individuals.

That said, a network approach offers a structural perspective that is currently missing from CDST research in the field of SLA and enables us to expand our research agenda beyond idiographic approaches, time-intensive approaches (Hiver et al., 2022). Taking steps toward generalizable findings, network analysis provides a means to quantitatively analyze the relationships between multiple variables and assess the relative importance of each variable within the system. Compared to other statistical techniques such as SEM, an advantage of network analysis is that it does not require *a priori* assumptions about unidirectional causal relations, but instead it allows for (hypothesized) bidirectional interactions between variables. As previously mentioned, other applications of network analysis such as the network comparison test make it possible for SLA researchers to test hypotheses and assess the extent to which systems can be generalized across different learner populations. Although network analysis is still relatively new, some researchers in clinical psychology have set out to examine its methodological validity and to determine the most appropriate metrics for assessing similarities between samples (Borsboom et al., 2017; Funkhouser et al., 2020). Researchers have also begun to assess the extent to which network analytic tools can inform the design of intervention studies. For example, Henry et al. (2020, p. 2) have developed a statistical testing procedure to assess the efficacy of an intervention, “determining if the dynamical systems of different people have the same optimal intervention studies.”

## Conclusion

In this article we provided a brief overview of research methods used by SLA researchers working within a CDST paradigm. We put forward network analysis as a way to model complex systems from a relation-intensive perspective and provided two examples of how to apply network analysis to two different datasets. In the first example we estimated a network model of L2 motivation, which provided a more fine-tuned picture of the potential relationships between motivational constructs compared to the original SEM analyses. In the second example we created a network model of individual differences in native language knowledge, showing how network analysis can model

the interconnectedness of individual difference constructs and different aspects of language knowledge.

While CDST researchers have made considerable advances in describing language development and changes in individual differences over time, the potential of relation-intensive approaches has not yet been explored. Through our two examples of network models, we hope to have illustrated that cross-sectional data does have a place in CDST research, and that network analysis is a useful technique to add to the CDST toolbox.

**Supplementary Materials.** To view supplementary material for this article, please visit <http://doi.org/10.1017/S0272263122000407>.

**Acknowledgments.** We would like to thank Han van der Maas, Wander Lowie, and the three anonymous reviewers for their invaluable feedback and suggestions on an earlier draft of this manuscript.

**Data Availability Statement.** The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at <https://osf.io/hjcvz/>

## References

- Al-Hoorie, A. H., & Hiver, P. (2022). Complexity theory: From metaphors to methodological advances. In A. H. Al-Hoorie & F. Szabó (Eds.), *Researching language learning motivation: A concise guide* (pp. 175–184). Bloomsbury Academic.
- American Psychological Association. (2017, February). Nonlinear methods for understanding complex dynamical phenomena in psychological science. <https://www.apa.org/science/about/psa/2017/02/dynamical-phenomena>
- Amerstorfer, C. M. (2020). The dynamism of strategic learning: Complexity theory in strategic L2 development. *Studies in Second Language Learning and Teaching*, 10, 21–44.
- Arndt, H. L., Granfeldt, J., & Gullberg, M. (2021). Reviewing the potential of the experience sampling method (ESM) for capturing second language exposure and use. *Second Language Research*. Advance online publication. <https://doi.org/10.1177/02676583211020055>
- Blanco, I., Contreras, A., Chaves, C., Lopez-Gomez, I., Hervas, G., & Vazquez, C. (2020). Positive interventions in depression change the structure of well-being and psychological symptoms: A network analysis. *The Journal of Positive Psychology*, 15, 623–628.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16, 5–13.
- Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91–121.
- Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., van Borkulo, C. D., van der Maas, H. L. J., & Cramer, A. O. J. (2017). False alarm? A comprehensive reanalysis of “evidence that psychopathology symptom networks have limited replicability” by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, 126, 989–999.
- Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks?, *Journal of Abnormal Psychology* 128, 892–903.
- Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, 8, e60188.
- Burger, J., Isvoranu, A. M., Lunansky, G., Haslbeck, J. M. B., Epskamp, S., Hoekstra, R. H. A., Fried, E. I., Borsboom, D., Blanken, T. F. (2022). Reporting standards for psychological network analyses in cross-sectional data. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000471>
- Bybee, J. L., & Beckner, C. (2009). Usage-based theory. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 1–26). Oxford University Press.
- Chang, P., & Zhang, L. J. (2021). A CDST perspective on variability in foreign language learners’ listening development. *Frontiers in Psychology*, 12, 1–17.
- Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, 178, 222–235.

- David, S. J., Marshall, A. J., Evanovich, E. K., & Mumma, G. H. (2018). Intraindividual dynamic network analysis—implications for clinical assessment. *Journal of Psychopathology and Behavioral Assessment*, 40, 235–248.
- De Bot, K. (2011). Researching second language development from a dynamic systems theory perspective. In M. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 5–24). John Benjamins.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10, 7–21.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.
- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. In N. Ellis & D. Larsen-Freeman (Eds.), *Language as a complex adaptive system* (pp. 230–248). Wiley-Blackwell.
- Dörnyei, Z. (2017). Conceptualizing L2 learner characteristics in a complex, dynamic world. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 79–96). John Benjamins.
- Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self images, sensory styles, and imagery capacity across two target languages. *Language Learning*, 63, 437–462.
- Dörnyei, Z., MacIntyre, P. D., & Henry, A. (2015). Introduction: Applying complex dynamic systems principles to empirical research on L2 motivation. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 1–7). Multilingual Matters.
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5, 155–174.
- Elliott, H., Jones, P. J., & Schmidt, U. (2020). Central symptoms predict posttreatment outcomes and clinical impairment in anorexia nervosa: A network analysis. *Clinical Psychology Science*, 8, 139–154.
- Ellis, N., & Wulff, S. (2018). Usage-based approaches to second language acquisition. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 37–56). John Benjamins.
- Epskamp, S. (2014). Network model selection using qgraph 1.3. <http://psychosystems.org/network-model-selection-using-qgraph-1-3-10/>
- Epskamp, S., & Fried, E. (2019). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23, 617–634.
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018a). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50, 195–212.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48, 1–18.
- Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 82, 904–927.
- Epskamp, S., Waldorp, L., Möttus, R., & Borsboom, D. (2018b). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, 53, 453–480.
- Evans, R., & Larsen-Freeman, D. (2020). Bifurcations and the emergence of L2 syntactic structures in a complex dynamic system. *Frontiers in Psychology*, 11, 1–12.
- Finnemann, A., Borsboom, D., Epskamp, S., & van der Maas, H. L. J. (2021). The theoretical and statistical Ising model: A practical guide in *R*. *Psych*, 3, 593–617.
- Fogal, G. (2022). Second language writing from a complex dynamic systems perspective. *Language Teaching*, 55, 193–210.
- Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry*, 31, 271–288.
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21, 1129–1164.
- Funkhouser, C. J., Correa, K. A., Gorka, S. M., Nelson, B. D., Phan, K. L., & Shankman, S. A. (2020). The replicability and generalizability of internalizing symptom networks across five samples. *Journal of Abnormal Psychology*, 129, 191–203.

- Gautier, R. (2019). Understanding socialisation and integration through social network analysis: American and Chinese students during a stay abroad. In M. Howard (Ed.), *Study abroad, second language acquisition and interculturality* (pp. 207–236). Multilingual Matters.
- Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE*, *12*, e0174035.
- Gregersen, T. (2020). Dynamic properties of language anxiety. *Studies in Second Language Learning and Teaching*, *10*, 67–87.
- Gregersen, T., Mercer, S., MacIntyre, P., Talbot, K., & Banga, C. A. (2020). Understanding language teacher wellbeing: An ESM study of daily stressors and uplifts. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688209658>
- Grice, J. W. (2004). Bridging the idiographic-nomothetic divide in ratings of self and others on the big five. *Journal of Personality*, *72*, 203–241.
- Guyon, H., Falissard, B., & Kop, J. (2017). Modeling psychological attributes in psychology—an epistemological discussion: Network analysis vs. latent variables. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.00798>
- Han, Z. (2019). Researching CDST: Promises and pitfalls. In Z. Han (Ed.), *Profiling learner language as a dynamic system* (pp. 156–166). Multilingual Matters.
- Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale (pp. 43–61). The Guilford Press.
- Haslbeck, J., Bringmann, L., & Waldorp, L. (2021). A tutorial on estimating time-varying vector autoregressive models. *Multivariate Behavioral Research*, *56*, 120–149.
- Henry, A. (2014). The dynamics of possible selves. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 83–94). Multilingual Matters.
- Henry, A. (2017). L2 motivation and multilingual identities. *The Modern Language Journal*, *101*, 548–565.
- Henry, T. R., Robinaugh, D., & Fried, E. I. (2020, April 1). On the control of psychological networks. *PsyArXiv*. <https://doi.org/10.31234/osf.io/7vpz2>
- Hevey, D. (2018). Network analysis: A brief overview and tutorial. *Health Psychology and Behavioral Medicine*, *6*, 301–328.
- Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology: Aligning theory and method. *Educational Psychologist*, *53*, 185–202.
- Hiver, P. (2014). Attractor states. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 20–28). Multilingual Matters.
- Hiver, P., & Al-Hoorie, A. H. (2016). A dynamic ensemble for second language research: Putting complexity theory into practice. *The Modern Language Journal*, *100*, 741–756.
- Hiver, P., & Al-Hoorie, A. H. (2020a). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of you, Dörnyei, and Csizér (2016). *Language Learning*, *70*, 48–102.
- Hiver, P., & Al-Hoorie, A. H. (2020b). *Research methods for complexity theory in applied linguistics*. Multilingual Matters.
- Hiver, P., Al-Hoorie, A. H., & Evans, R. (2022). Complex dynamic systems theory in language learning: A scoping review of 25 years of research. *Studies in Second Language Acquisition*, *44*, 913–941.
- Hiver, P., & Larsen-Freeman, D. (2019). Motivation: It is a relational system. In A. Al-Hoorie & P. MacIntyre (Eds.), *Contemporary language motivation theory* (pp. 285–303). Multilingual Matters.
- Hiver, P., & Papi, M. (2019). Complexity theory and L2 motivation. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 117–137). Palgrave Macmillan.
- Hulstijn, J. (2020). Proximate and ultimate explanations of individual differences in language use and language acquisition. *Dutch Journal of Applied Linguistics*, *9*, 21–37.
- Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching*, *10*, 89–109.
- Khajavy, G. H., MacIntyre, P. D., Taherian, T., & Ross, J. (2021). Examining the dynamic relationships between willingness to communicate, anxiety and enjoyment using the experience sampling method. In N. Zarrinabadi, & M. Pawlak (Eds.), *New perspectives on willingness to communicate in a second language* (pp. 169–197). Springer.
- Kievit, R. A. (2020). Sensitive periods in cognitive development: A mutualistic perspective. *Current Opinion in Behavioral Sciences*, *36*, 144–149.

- Kliesch, K., & Pfenninger, S. (2021). Cognitive and socioaffective predictors of L2 microdevelopment in late adulthood: A longitudinal intervention study. *The Modern Language Journal*, 105, 237–266.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18, 141–165.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27, 590–619.
- Larsen-Freeman, D. (2017). Complexity theory: The lessons continue. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 1–10). John Benjamins.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Larsen-Freeman, D., Schmid, M., & Lowie, W. (2011). Introduction: From structure to chaos. Twenty years of modeling bilingualism. In M. Schmid & W. Lowie (Eds.), *Modeling Bilingualism: From structure to chaos. In honor of Kees de Bot* (pp. 1–12). John Benjamins.
- Li, C., Dewaele, J., & Jiang, G. (2020). The complex relationship between classroom emotions and EFL achievement in China. *Applied Linguistics Review*, 11, 485–510.
- Lowie, W., van Dijk, M., Chan, H., & Verspoor, M. (2017). Finding the key to successful L2 learning in groups and individuals. *Studies in Second Language Learning and Teaching*, 7, 127–148.
- Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, 69, 184–206.
- Lutz, W., Schwartz, B., Hofmann, S. G., Fisher, A. J., Husen, K., & Rubel, J. A. (2018). Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: A methodological proof-of-concept study. *Scientific Reports*, 8, 7819.
- MacIntyre, P. (2012). The idiodynamic method: A closer look at the dynamics of communication traits. *Communication Reports*, 29, 361–367.
- MacIntyre, P. (2020). Expanding the theoretical base for the dynamics of willingness to communicate. *Studies in Second Language Learning and Teaching*, 10, 111–131.
- MacIntyre, P. D., MacKay, E., Ross, J., & Abel, E. (2017). The emerging need for methods appropriate to study dynamic systems. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 97–122). John Benjamins.
- Marchand, G. C., & Hilpert, J. C. (2018). Design considerations for education scholars interested in complex systems research. *Complicity: An International Journal of Complexity and Education*, 15, 31–44.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41, 954–969.
- Mercer, S. (2011a). The self as a complex dynamic system. *Studies in Second Language Learning and Teaching*, 1, 57–82.
- Mercer, S. (2011b). Understanding self-concept in the FLL context. In S. Mercer, *Towards an understanding of language learner self-concept* (pp. 35–71). Springer.
- Mercer, S. (2014). Social network analysis and complex dynamic systems. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 83–94). Multilingual Matters.
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2, 201–218.
- Nitta, R., & Baba, K. (2018). Understanding benefits of repetition from a complex dynamic systems perspective. In M. Bygate (Ed.), *Language learning through task repetition* (pp. 279–309). John Benjamins.
- O'Malley, A., & Onella, J. (2019). Introduction to social network analysis. In A. Levy, S. Goring, C. Gatsonis, B. Sobolev, E. van Ginneken, & R. Busse (Eds.), *Health services evaluation* (pp. 617–660). Springer.
- Palloti, G. (2022). Cratylus' silence: On the philosophy and methodology of complex dynamic systems theory in SLA. *Second Language Research*, 38, 689–701.
- Papi, M., & Hiver, P. (2020). Language learning motivation as a complex dynamic system: A global perspective of truth, control, and value. *The Modern Language Journal*, 104, 209–232.
- Paradowski, M. B., Jarynowski, A., Jelińska, M., & Czopek, K. (2021). Selected poster presentations from the American Association of Applied Linguistics conference, Denver, USA, March 2020. *Language Teaching*, 54, 139–143.
- Pfenninger, S. E. (2020). The dynamic multicausality of age of first bilingual language exposure: Evidence from a longitudinal content and language integrated learning study with dense time serial measurements. *The Modern Language Journal*, 104, 662–686.

- Piniel, K., & Cszér, K. (2014). Changes in motivation, anxiety and self-efficacy during the course of an academic writing seminar. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 164–194). Multilingual Matters.
- R Core Team (2020). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reifegerste, J. (2021). The effects of aging on bilingual language: What changes, what doesn't, and why. *Bilingualism: Language and Cognition*, *24*, 1–17.
- Rodrigues, F. (2019). Network centrality: An introduction. In E. Macau (Ed.), *A mathematical modeling approach from nonlinear dynamics to complex systems. nonlinear systems and complexity* (pp. 177–196). Springer.
- Rouquette, A., Pingault, J., Fried, E. I., Orri, M., Falissard, B., Kossakowski, J. J., ... Borsboom, D. (2018). Emotional and behavioral symptom network structure in elementary school girls and association with anxiety disorders and depression in adolescence and early adulthood: A network analysis. *JAMA Psychiatry*, *75*, 1173–1181.
- Sachisthal, M. S. M., Jansen, B. R. J., Peetsma, T. T. D., Dalege, J., van der Maas, Han L. J., & Raijmakers, M. E. J. (2019). Introducing a science interest network model to reveal country differences. *Journal of Educational Psychology*, *111*, 1063–1080.
- Saito, K., Macmillan, K., Mai, T., Suzukida, Y., Sun, H., Magne, V., Ilkan, M., & Murakami, A. (2020). Developing, analyzing and sharing multivariate datasets: Individual differences in L2 learning revisited. *Annual Review of Applied Linguistics*, *40*, 9–25.
- Salvatore, S., & Valsiner, J. (2010). Between the general and the unique: Overcoming the nomothetic versus idiographic opposition. *Theory & Psychology*, *20*, 817–833.
- Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *The Modern Language Journal*, *101*, 369–390.
- Spearman, C. (1904). “General intelligence,” objectively determined and measured. *The American Journal of Psychology*, *15*, 201–292.
- Spencer, J. P., & Schöner, G. (2003). Bridging the representational gap in the dynamic systems approach to development. *Developmental Science*, *6*, 392–412.
- Tan, K. H., & Shojamanesh, V. (2019). Usage-based and universal grammar-based approaches to second language acquisition. In O. Lutsenko & G. Lutsenko (Eds.), *Active Learning—Theory and Practice*. Intech Open. <https://doi.org/10.5772/intechopen.89737>
- Tiv, M., Gullifer, J. W., Feng, R. Y., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics*, *56*, 1–17.
- van Bork, R., Rhemtulla, M., Waldorp, L. J., Kruijs, J., Rezvanifar, S., & Borsboom, D. (2019). Latent variable models and networks: Statistical equivalence and testability. *Multivariate Behavioral Research*, *56*, 175–198.
- van Bork, R., van Borkulo, C. D., Waldorp, L. J., Cramer, A. O. J., & Borsboom, D. (2018). Network models for clinical psychology. In J. T. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed.). Wiley. <https://doi.org/10.1002/9781119170174.epcn518>
- van Borkulo, C. D., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000476>
- van der Maas, Han L. J., Dolan, C. V., Grasman, Raoul P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861.
- van der Maas, Han L. J., Kan, K., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, *5*, 16.
- van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*, 3–53.
- van Geert, P. (2019). Dynamic systems, process and development. *Human Development*, *63*, 153–179.
- van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, *25*, 408–442.
- van Geert, P., & van Dijk, M. (2021). Thirty years of focus on individual variability and the dynamics of processes. *Theory & Psychology*, *31*, 405–410.

- Vu, T., Magis-Weinberg, L., Jansen, B. R. J., van Atteveldt, N., Janssen, T. W. P., Lee, N. C., van der Maas, H. L. J., Raijmakers, M. E. J., Sachisthal, M. S. M., & Meeter, M. (2021). Motivation-achievement cycles in learning: A literature review and research agenda. *Educational Psychology Review*, 34, 39–71.
- Waninge, F., Dörnyei, Z., & de Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, 98, 704–723.
- You, C., Dörnyei, Z., & Csizér, K. (2016). Motivation, vision, and gender: A survey of learners of English in China. *Language Learning*, 66, 94–123.
- Zappa-Hollman, S., & Duff, P. A. (2014). Academic English socialization through individual networks of practice. *TESOL Quarterly*, 49, 333–368.

---

**Cite this article:** Freeborn, L., Andringa, S., Lunansky, G. and Rispen, J. (2023). Network analysis for modeling complex systems in SLA research. *Studies in Second Language Acquisition*, 45, 526–557. <https://doi.org/10.1017/S0272263122000407>