

### Research Article

## Emerging Themes and Research Trends in Identity and Language Studies: A Bibliometric Analysis

Sahar Zahed Alavi 

Assistant Professor, Department of English Language and Literature, University of Bojnord, Bojnord, Iran,  
[s.zahedalavi@ub.ac.ir](mailto:s.zahedalavi@ub.ac.ir)

#### ARTICLE INFO

##### *Submission History*

Received: 2026-02-13

Accepted: 2026-04-01

##### *Keywords*

Keyword co-occurrence  
Overlay visualization  
Co-citation of references

##### *How to cite this article*

Alavi, S. Zahed (2026). Emerging Themes and Research Trends in Identity and Language Studies: A Bibliometric Analysis. *Language, Identity and the Digital Realm*, 2(2), 1-19. doi: 10.22034/lidr.2026.2085376.1040

#### ABSTRACT

Research on language and identity has expanded considerably over the last two decades. Despite this expansion, the field remains theoretically diverse and methodologically fragmented. The present study provides a bibliometric review to map the intellectual structure, research trends, and scholarly influence within identity and language studies. Using bibliometric techniques, the study explores the relationship between author productivity and scholarly impact. Keyword co-occurrence and overlay visualization analyses are used to identify major thematic clusters and distinguish well-established topics from emerging or underexplored research areas. In addition, citation analysis is conducted to identify the most influential papers, while reference co-citation analysis is used to reveal the theoretical and conceptual foundations underpinning the field. The findings show that the higher number of papers does not necessarily lead to higher research impact. The keyword co-occurrence analysis reveals that identity and language research is organized around five interconnected thematic clusters, including sociolinguistic and ideological perspectives, bilingualism and heritage language, educational and professional identity, discourse and methodology-orientation, and psychological-affective dimensions of identity. As the overlay visualization shows, the emergence of topics such as teacher identity, professional identity, and agency reflects a growing concern. In addition, Norton (1995) and Heller (2003) are the most influential theories in the field. This study contributes to a deeper understanding of research trajectories and offers insights to guide future investigations on language and identity.



## Introduction

Research on language learning has experienced a remarkable expansion over the last two decades, driven by intensified globalization, large-scale migration, and the widespread use of digital communication technologies (Bose et al., 2023). This growth underscores the increasing social (Ushioda, 2017), educational (Hofer, 2025), and economic (Rozenvalde & Martena, 2024) importance of multilingual competence in contemporary societies. Alongside this expansion, scholars have made substantial progress in developing theoretical and empirical investigations that explain how language learning is shaped by social context (Liu & Chao, 2017), learner agency (Moradi, 2020), and interactional environments (Yung, 2025). These developments reflect an increasingly sophisticated and integrative understanding of the language learning process. Within this broad landscape, one prominent direction of inquiry has focused on the ways in which learners' identities are implicated in and shaped by language learning experiences, signaling a shift from general language learning research toward a more specialized scholarly domain (Perlman, 2020).

Within this specialized domain, the intersection of identity and language learning has emerged as a distinct and interdisciplinary area of investigation. Prior research has identified a range of well-established themes and methodological practices, including identity negotiation in classroom interaction (Fisher et al., 2018), multilingual contexts (Abourehab & Azaz, 2020), and study-abroad experiences (Li, 2025). Scholars have also drawn on influential theoretical frameworks such as poststructuralist identity theory (e.g., Fisher et al., 2018), investment theory (e.g., Sung, 2019), and communities of practice

(e.g., Evnitskaya & Morton, 2011) to examine how identity formation and language learning mutually inform one another.

However, despite these advances, there remains a clear lack of comprehensive and systematic evidence on identity and language learning. Most traditional review studies are limited in scope, relying on small samples, restricted databases, short time spans, or narrative and interpretive synthesis methods (Dixon-Woods et al., 2005). Furthermore, little attention has been paid to large-scale quantitative mapping of research on identity and language. Consequently, it remains uncertain how the knowledge base has been constructed over time and which directions are truly dominant or emerging (Larsen-Freeman, 2018).

This uncertainty has tangible negative consequences for both research and practice. When the development and structure of the literature are not clearly understood, scholars may struggle to position new studies effectively, identify genuine research needs, or avoid unnecessary duplications (McKeown & Mir, 2021). In addition, the lack of systematic evidence hinders efficient allocation of academic resources, thereby preventing funding bodies from prioritizing high-impact research areas (Dennison et al., 2025). Given the growth of publications in this domain, the need to address this problem and adopt a more rigorous and comprehensive approach to synthesize existing research has become increasingly urgent.

Thus, the present study aims to provide a bibliometric review of research on identity and language by systematically examining its thematic and developmental trajectories. Specifically, this study investigates the possible relationship between authors' scholarly productivity and their academic impact,

thereby shedding light on patterns of knowledge production within the field. It also traces the evolution of research output on language and identity to identify growth trends and shifts in scholarly attention. Furthermore, the study explores the major thematic clusters that characterize identity and language research and examines the conceptual relationships among these themes. Moreover, this study distinguishes between well-established research topics and emerging or underexplored areas, providing insight into potential directions for future research. Finally, it examines the most influential documents in the field and uncovers the theoretical and conceptual foundations underpinning identity and language research.

## Literature Review

The relationship between language and identity has evolved into one of the most theoretically complex and interdisciplinary domains within sociolinguistics, applied linguistics, and linguistic anthropology (Nie, 2025). Early studies conceptualized identity as a stable and predetermined attribute reflected through language use (Tong et al., 1999). However, contemporary research has fundamentally challenged this essentialist orientation by advancing constructivist and poststructuralist paradigms that conceptualize identity as socially constructed, fluid, and discursively negotiated (Sultana, 2022). While this paradigmatic shift has significantly broadened analytical perspectives, it has also produced theoretical fragmentation, as scholars increasingly adopt diverse frameworks that conceptualize identity from competing epistemological standpoints (Kaplan & Garner, 2017).

One of the most influential contributions to identity scholarship is Norton's (1995; 2015)

investment theory, which reconceptualizes language learning as a socially mediated process shaped by learners' access to power, symbolic capital, and imagined future identities. The investment theory effectively challenges traditional motivation-based frameworks that emphasize individual psychological dispositions, instead foregrounding sociopolitical structures and learner positioning (Darvin & Norton, 2015). Nevertheless, despite its conceptual strength, the theory has been critiqued for its tendency to privilege learner agency while insufficiently theorizing structural inequalities that restrict access to linguistic resources. This imbalance reflects a broader tension in identity research between recognizing individuals' agentic identity performances and acknowledging institutional constraints that regulate participation and belonging (Block, 2013).

Sociocultural perspectives, including language socialization and community-of-practice frameworks, further extend identity research by conceptualizing language learning as participation in socially organized practices through which individuals simultaneously acquire linguistic competence and social membership (Kinginger, 2011). These frameworks have been particularly valuable in demonstrating the interactional processes through which identities are co-constructed. However, scholars have questioned their capacity to account adequately for macro-level ideological forces and power hierarchies that shape participation opportunities (Mori, 2014). Consequently, while sociocultural models provide detailed insights into localized identity negotiations, they risk underestimating structural inequalities embedded in broader sociopolitical systems (Berman & Smith, 2021).

Poststructuralist approaches have attempted to address these limitations by situating identity formation within complex networks of power, discourse, and ideology (Zembylas, 2003). By rejecting binary conceptualizations such as stability versus fluidity or agency versus structure, poststructuralists reconceptualize identity as inherently relational (Schiller, 2012). The concept of indexicality further enriches this perspective by illustrating how linguistic features signal social meanings and position speakers within stratified social structures (Li & Roberts, 2023). Despite their analytical sophistication, poststructuralist frameworks have been criticized for theoretical abstraction and limited empirical operationalization, which complicates efforts to synthesize findings across studies (Block, 2007). This limitation has contributed to a proliferation of theoretically diverse but empirically disconnected research strands within language and identity scholarship (Pavlenko & Blackledge, 2004).

Language ideology research has further highlighted the role of sociocultural belief systems in regulating linguistic legitimacy and shaping identity construction (Vélez-Rendón, 2010). Some studies demonstrate how dominant monolingual ideologies marginalize multilingual practices and reinforce hierarchical social classifications (Kiramba, 2018; Sah, 2024). However, integrating micro-level processes with macro-level discourse analysis highlights how individuals negotiate and reinterpret ideological constraints (Humphrey et al., 2019).

Research on multilingualism has further complicated traditional conceptualizations of identity by emphasizing linguistic hybridity and the strategic deployment of multilingual repertoires (Benzehaf, 2021; Han, 2025).

Multilingual speakers frequently engage in complex identity negotiations that transcend conventional sociolinguistic categorizations, thereby challenging monolingual bias in language research (Henry, 2017). Translanguaging theory has emerged as a transformative framework that reconceptualizes multilingual communication as an integrated and dynamic process rather than alternation between discrete linguistic systems. This perspective has significantly advanced the understanding of multilingual identity expression by highlighting speakers' creative agency in deploying linguistic resources (Wei, 2017).

Despite the substantial diversification of theoretical perspectives and empirical contexts, language and identity research remains conceptually fragmented and methodologically dispersed. Furthermore, there is a notable absence of large-scale analyses examining patterns of scholarly productivity, citation influence, and conceptual clustering within identity and language research.

Given these limitations, the bibliometric analysis offers a valuable methodological framework for systematically mapping the development of language and identity scholarship. By examining publication trends, citation networks, keyword co-occurrence patterns, and co-citation relationships, a bibliometric review enables researchers to identify dominant theoretical traditions, established and emerging research themes, and intellectual linkages across disciplines. Therefore, the following research questions are raised:

1. How has research output on *language and identity* evolved over the past two decades?
2. Is there any relationship between the authors' productivity and their impact?

3. What are the major thematic clusters in research on *identity and language*, and how are these themes conceptually related?
4. Which research topics in *identity and language* are well-established, and which represent emerging or underexplored areas?
5. What are the most influential documents in the field of *identity and language* research?
6. What theoretical and conceptual foundations underpin research on *identity and language*, and how are these theoretical traditions interconnected?

## Methods

### Research Design

This study adopts a bibliometric research design to systematically map the intellectual structure, thematic evolution, and research trends on identity and language. Bibliometric methods enable large-scale, replicable analyses of scientific publications. They are particularly well-suited for identifying research trends, influential works, and conceptual linkages within an interdisciplinary field such as language and identity studies (Pessin et al., 2022).

### Data Source and Search Strategy

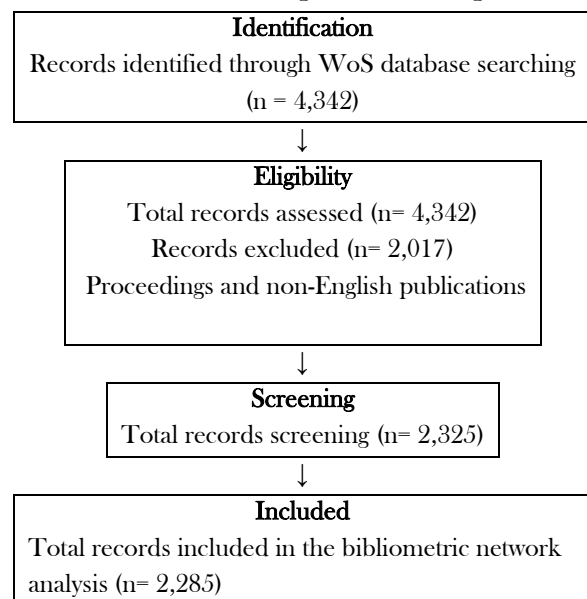
The bibliographic data were retrieved from the Web of Science Core Collection, including Science Citation Index Expanded, Social Sciences Citation Index, Arts and Humanities Citation Index, and Emerging Sources Citation Index. A topic search was conducted using the combination of the terms identity and language (linguistic) applied to titles. The search strategy was TI=((language OR linguistic) AND identity) AND (Article OR Book Chapters OR Review Article) (Document Types) AND English

(Languages). The timespan was not restricted. The search link was <https://www.webofscience.com/wos/woscc/su/mmary/2f09b9e9-1c29-41c9-8afe-69d9215bdfc2-019e9ef962/relevance/1>.

4,342 papers were exported from the Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (A&HCI), and Emerging Sources Citation Index (ESCI). After refining the search and excluding proceeding papers and non-English publications, and removing duplicates, 2,285 documents remained for investigation. The PRISMA chart is presented in Figure 1.

**Figure 1**

*PRISMA chart showing the screening of data*



### Bibliometric Analysis Techniques

To examine the research questions and to offer a multidimensional understanding of trends and patterns in identity and language research, multiple bibliometric techniques were employed using VOSviewer software. First, to examine the evolution of research output over time (the first research question), a graph showing the annual publication counts and growth trends was retrieved from the WoS

search report. It provided insights into the developmental trajectory of identity and language research and helped identify periods of accelerated scholarly attention. Second, to examine the relationship between the number of papers of an author and the number of citations of the papers (the second research question), citation analysis of authors was conducted. Third, to identify the major thematic clusters within the field (the third research question), keyword co-occurrence analysis was conducted. By visualizing keyword networks, the analysis revealed dominant research themes and the conceptual relationships among them. Fourth, to distinguish between established and emerging research topics (the fourth research question), overlay visualization was used. In this analysis, keywords are colored according to their average publication year, allowing newer topics to be differentiated from long-standing themes. This temporal dimension provided

insights into the shifting priorities and emerging directions within identity and language research. Fifth, to identify the most influential documents in the field (the fifth research question), citation analysis of documents was performed. Sixth, to uncover the theoretical and conceptual foundations of the field (the sixth research question), reference co-citation analysis was performed. This method examines how frequently pairs of references are cited together, thereby revealing clusters of foundational theories and intellectual traditions.

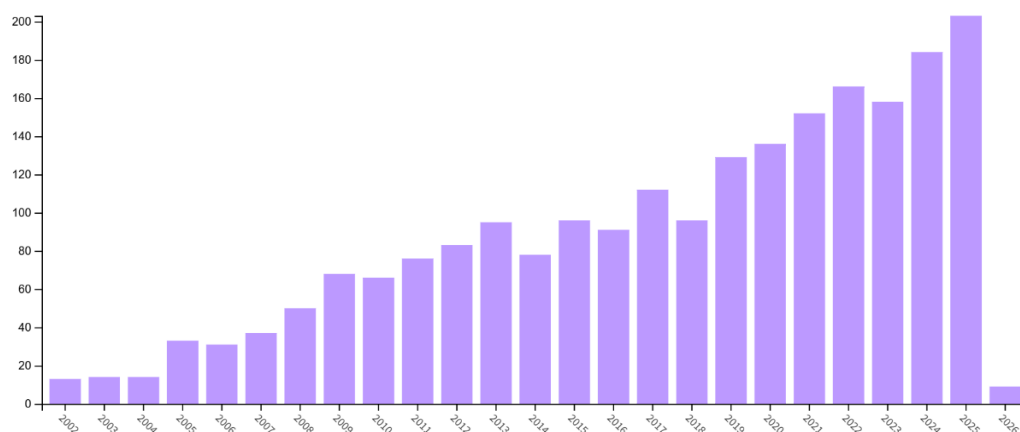
## Results

### The trend in annual publications on language and identity

Figure 2 illustrates the annual number of publications on language and identity indexed in Web of Science from 2002 to 2025, revealing a steady increase in research output.

**Figure 2**

*The number of publications on language and identity indexed in Web of Science from 2002 to 2025*



### The relationship between the authors' productivity and their impact

To examine the relationship between authors' productivity and citation impact in identity and language research, the citation

analysis of authors was conducted. Table 1 presents the ten most productive authors, their citations, and their total link strength. The table indicates that a higher number of publications does not necessarily correspond

to a higher number of citations. For instance, Mostafa Nazari, who had the highest number of publications (16 documents), accumulated 275 citations, which is lower than the citation counts of Bonny Norton (395 citations) and Yongcan Liu (392 citations), despite both authors producing fewer documents (9 and 6 documents, respectively). Similarly, Juyoung Song produced only five documents but received a relatively high citation count (308 citations), further suggesting that limited publication output can still generate substantial scholarly influence. In addition, authors with moderate productivity levels demonstrated inconsistent citation outcomes. For instance, Sedigheh Karimpour produced six documents

and received 168 citations, whereas Rui Yuan and John Trent, with five and seven documents, received 122 and 123 citations, respectively. These findings reinforce the observation that publication quantity alone does not predict citation influence.

Furthermore, total link strength, which reflects collaborative and intellectual connectivity, appears to complement citation performance. Authors such as Nazari, Song, and Karimpour demonstrated relatively high link strength values, suggesting that collaborative engagement may partially contribute to scholarly visibility, although it does not consistently guarantee higher citation counts.

**Table 1**

*The ten most productive authors and their citations according to the citation analysis of authors*

Author	Documents	Citations	Total link strength
Nazari, Mostafa	16	275	24
Norton, Bonny	9	395	7
Park, Mi Yung	8	52	1
Trent, John	7	123	3
Karimpour, Sedigheh	6	168	17
Liu, Yongcan	6	392	3
Song, Juyoung	5	308	18
Sung, Chit Cheung Matthew	5	88	0
Weng, Zhenjie	5	51	6
Yuan, Rui	5	122	14

### **The major thematic clusters in research on *identity and language***

Figure 3 presents the keyword co-occurrence network of research on identity and language generated using VOSviewer. The map visualizes the field's conceptual structure based on the frequency and co-occurrence of author keywords. Node size reflects keyword frequency, link strength represents the degree of co-occurrence, and colors indicate distinct thematic clusters.

At the center of the network, the keywords *identity* and *language* appear as the most prominent and strongly connected nodes,

confirming their foundational role in structuring the field. Surrounding these core concepts, five well-defined thematic clusters can be identified. The green cluster is primarily associated with sociolinguistic and ideological perspectives, including keywords such as *language ideology*, *language policy*, *language choice*, *translanguaging*, *indexicality*, *sociolinguistics*, and *migration*. This cluster illustrates research examining how identity is shaped through language practices within broader sociopolitical, ideological, and migratory contexts. The blue cluster focuses on bilingualism, heritage language, and youth



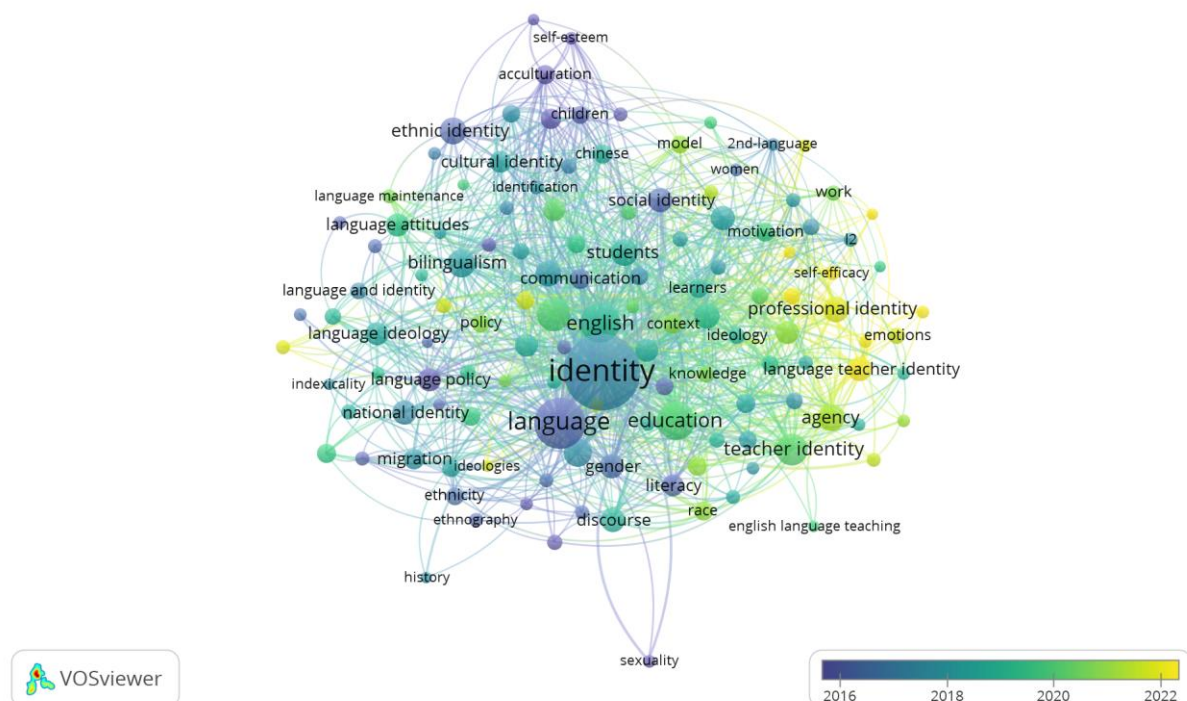
social identity, culture, ethnography, and discourse) appear in earlier publication years and occupy central, highly connected positions, indicating well-established research areas. The prominence and early coloration of these keywords suggest that identity research initially developed within sociocultural, ethnographic, and discourse-oriented traditions, focusing on identity as socially constructed, historically situated, and discursively negotiated. Their strong co-occurrence links indicate that these themes have served as theoretical foundations for subsequent work. In addition, keywords such as bilingualism and multilingualism appear as relatively established topics, reflecting

sustained scholarly interest in identity formation in linguistically diverse contexts.

More recent keywords (green to yellow nodes) indicate emerging or expanding areas of inquiry. Notably, these areas include teacher identity, professional identity, agency, English language teaching, teachers, students, and education. These keywords suggest a shift toward practice-oriented and professionally grounded research, particularly within educational and English language teaching contexts. The increasing prominence of teacher identity and agency indicates growing attention to educators' roles as active agents who negotiate institutional constraints, pedagogical beliefs, and policy demands.

**Figure 4**

*Overlay visualization of keyword co-occurrence*



### The most influential documents

To identify the most influential documents on identity and language, a citation analysis of documents was conducted. Table 2 presents the top-cited documents in the dataset. The

analysis reveals that Peirce (1995) is the most influential document in the field, with 1,128 citations and 10 citation links, indicating both high scholarly impact and strong interconnection with other influential works.

This study stands out as the central reference point in identity and language research. The second most cited document is Heller (2003), which received 591 citations and showed two citation links, suggesting considerable influence in sociolinguistic and identity-related scholarship. Similarly, Bucholtz (1999) gained 408 citations. Other highly cited works include Dunn and Andrews (2015) with 376 citations, Morita (2004) with 353 citations, and Phinney et al. (2001) with 338 citations. These studies represent diverse research strands, including identity development, second language learning, and intercultural identity formation. Moderately influential publications include

Huffaker and Calvert (2005) with 296 citations, Giles and Johnson (1987) with 283 citations, and Fiol (2002) with 273 citations. Although these studies demonstrate strong citation counts, their relatively limited citation links suggest more specialized or context-specific scholarly influence. Finally, Kanno and Stuart (2011), with 246 citations and four citation links, represent a significant contribution to research on teacher identity and second language teacher development, highlighting growing scholarly attention to professional identity within language education.

**Table 2**

*The ten most influential papers according to the citation analysis of documents*

Document	Citations	Links
Norton, B. (1995). Social identity, investment, and language learning. <i>TESOL Quarterly</i> , 29(1), 9-31.	1128	10
Heller, M. (2003). Globalization, the new economy, and the commodification of language and identity. <i>Journal Of Sociolinguistics</i> , 7(4), 473-492.	591	2
Bucholtz, M. (1999). Why be normal: Language and identity practices in a community of nerd girls. <i>Language in Society</i> , 28(2), 203-223.	408	0
Dunn, D. S., & Andrews, E. E. (2015). Person-first and identity-first language: Developing psychologists' cultural competence using disability language. <i>American Psychologist</i> , 70(3), 255-264.	376	2
Morita, N. (2004). Negotiating participation and identity in second language academic communities. <i>TESOL Quarterly</i> , 38(4), 573-603.	353	0
Phinney, J., Romero, I., Nava, M., & Huang, D. (2001). The role of language, parents, and peers in ethnic identity among adolescents in immigrant families. <i>Journal of Youth and Adolescence</i> , 30(2), 135-153.	338	4
Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. <i>Journal of computer-mediated communication</i> , 10(2), JCMC10211.	296	0
Giles, H., & Johnson, P. (1987). Ethnolinguistic identity theory: A social psychological approach to language maintenance. <i>International Journal of the Sociology of Language</i> , (68), 69-99.	283	1
Fiol, C. M. (2002). Capitalizing on paradox: The role of language in transforming organizational identities. <i>Organization Science</i> , 13(6), 653-666.	273	0
Kanno, Y., & Stuart, C. (2011). Learning to become a second language teacher: Identities-in-practice. <i>Modern Language Journal</i> , 95(2), 236-252.	246	4

### Theoretical and conceptual foundations underpin research on identity and language

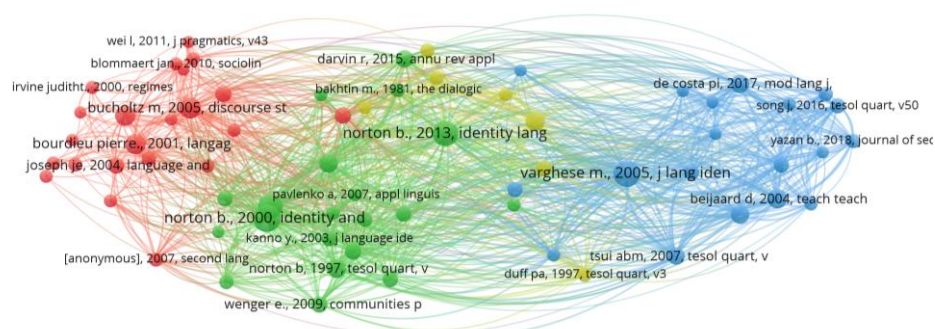
To identify the theoretical and conceptual foundations underpinning the research on identity and language and to examine how these foundations are interconnected, a co-

citation analysis was conducted. The resulting network (see Figure 5) reveals four theoretically coherent clusters, each representing a major intellectual tradition that has shaped the field.

The red cluster is grounded in Sociolinguistic and Ideological Perspectives on Identity, with influential works such as Bourdieu (2001), Blommaert (2010), Bucholtz and Hall (2005), Joseph (2004), and Irvine and Gal (2000). These studies conceptualize identity as socially constructed through discourse, ideology, and relations of power. The dense co-citation links within this cluster indicate a strong and stable theoretical foundation that foregrounds macro-level social processes in identity construction. The green cluster centers on learner identity and investment in language learning, with key contributions from Norton (1997, 2000, 2013), Wenger (2009), Pavlenko (2007), and Kanno (2003). This tradition views identity as dynamic, negotiated through participation in social practices, and closely tied to learners' investment and agency. The central location of this cluster indicates its pivotal role in linking broader sociolinguistic theory with empirical research in language learning contexts. The yellow cluster is anchored by Bakhtin's (1981) theory of dialogism, alongside works such as Darwin and Norton (2015), which integrate dialogic theory with contemporary identity frameworks. This cluster provides the dialogic and poststructural models of identity for understanding identity as relational, multi-voiced, and constantly emerging through

interaction. The strong co-citation links between the yellow cluster and both the green and red clusters indicate that dialogism functions as a conceptual hinge, connecting macro-level sociolinguistic concerns with micro-level processes of meaning-making and interaction. Rather than forming an isolated tradition, dialogic theory permeates identity research by offering a unifying lens through which discourse, agency, and social positioning are theorized. In addition, the blue cluster comprises research on identity in educational and professional settings, particularly language teacher identity. Seminal works include Varghese et al. (2005), Beijaard et al. (2004), Tsui (2007), Duff and Uchida (1997), and De Costa and Norton (2017). This cluster emphasizes identity as shaped by institutional constraints, policy, emotion, and professional practice, drawing heavily on sociocultural and dialogic theories.

In sum, several references, most notably Bakhtin (1981) and Norton (2013), occupy central positions with strong cross-cluster connections, highlighting their integrative role across theoretical traditions. The co-citation structure demonstrates that identity and language research is underpinned by multiple interdependent theoretical foundations, rather than discrete or competing frameworks.

**Figure 5***The co-citation analysis of references*

## Discussion

### The relationship between the authors' productivity and their impact

The present findings suggest that research productivity and citation impact are not strongly related within identity and language research. Although highly productive authors tend to contribute significantly to knowledge production, citation influence appears to depend more on the theoretical, conceptual, and methodological significance of individual publications rather than publication number alone. The prominence of scholars such as Bonny Norton illustrates this pattern. Norton's relatively limited number of publications, combined with exceptionally high citation counts, highlights the enduring theoretical influence of foundational works. This supports previous research indicating that seminal theoretical contributions often generate sustained citation impact regardless of publication frequency (Zhang & Glänzel, 2012). Norton's (1995, 2015) work, particularly in establishing the concept of investment and linking identity to language learning and social power, has become central to identity research, which may explain its broad scholarly uptake.

Similarly, Liu's high citation performance with fewer publications suggests that research addressing contemporary issues such as language teacher identity (Liu & Yueting, 2011) and multilingual identity construction and globalization resonates (Fisher et al., 2018) strongly with current scholarly interests. These findings indicate that research relevance and conceptual innovation may be stronger predictors of citation impact than research productivity alone (Jamali et al., 2023).

Conversely, authors with high publication output but comparatively lower citation counts, such as Nazari, demonstrate that productivity contributes to visibility and collaborative integration but does not necessarily translate into proportionally higher citation impact. This suggests that sustained productivity may facilitate academic presence and network formation, yet citation recognition remains closely tied to the perceived significance of individual contributions (Mazzucchelli et al., 2019). In addition, the role of collaborative connectivity, reflected in total link strength, suggests that scholarly networking and research collaboration may enhance knowledge

dissemination. However, collaboration alone does not guarantee citation impact, indicating that the intellectual contribution of research remains a critical factor (Jamali et al., 2023).

### **The major thematic clusters in research on identity and language**

The keyword co-occurrence analysis reveals that identity and language research is organized around multiple, interconnected thematic traditions, rather than a single dominant paradigm. The central positioning of identity and language confirms that identity is not treated as an ancillary concept, but as a core analytical lens through which language-related phenomena are examined.

Figure 2 reflects several prominent findings. The first one is the strong presence of a sociolinguistic-ideological cluster, which foregrounds concepts such as language ideology, policy, indexicality, and migration. This reflects a long-standing tradition in sociolinguistics that views identity as socially constructed, politically situated, and embedded in power relations. The prominence of this cluster suggests sustained scholarly attention to macro-level structures shaping identity, particularly in contexts of globalization and mobility (e.g., Blommaert, 2010; Dennison et al., 2025; Heller, 2003). Second, equally significant is the bilingualism and heritage language cluster, which highlights the central role of multilingual practices in identity construction. The strong links between bilingualism, heritage language, and youth indicate that identity research has increasingly focused on multilingual speakers navigating linguistic, cultural, and ethnic affiliations across contexts. This aligns with broader shifts toward transnational and post-national understandings of identity (e.g., Hofer, 2025; Rozenvalde & Martena, 2024). Third, the educational and professional

identity cluster, particularly the dense concentration around language teacher identity, points to the growing influence of identity-oriented approaches in applied linguistics and language education (e.g., Varghese et al., 2005; De Costa & Norton, 2017; Zembylas, 2003). The co-occurrence of agency, power, and narrative inquiry suggests that teacher identity research is deeply informed by poststructuralist and sociocultural perspectives, emphasizing identity as dynamic, negotiated, and context-dependent. Fourth, the emergence of a psychological-affective cluster linking motivation, investment, self-efficacy, and emotions indicates an important theoretical integration between identity research and affective dimensions of language learning (e.g., Henry, 2017; Ushioda, 2017). This convergence reflects a move beyond purely cognitive models toward more holistic accounts of language learning that recognize learners' emotional experiences and social positioning. Fifth, the discourse and methodology-oriented cluster underscores the central role of qualitative and interpretive methods (e.g., discourse analysis and ethnography) in identity research. The visibility of these methods highlights the field's commitment to capturing identity as lived, performed, and narratively constructed (e.g., Bucholtz & Hall, 2005; Pessin et al., 2020).

Overall, the network reveals a densely interconnected structure, indicating that research on identity and language is highly interdisciplinary, with strong conceptual linkages across sociolinguistic, educational, discursive, and psychological perspectives. The dense interconnections across clusters indicate not fragmentation, but conceptual convergence, positioning identity as a unifying construct across diverse strands of language research.

### **Established and emerging research topics based on overlay visualization**

The overlay visualization reveals a clear diachronic progression in identity and language research, moving from macro-level theoretical concerns toward micro-level, context-sensitive investigations. Early research, grounded in discourse analysis, ethnography, and sociocultural theory, established identity as a dynamic and socially constructed phenomenon. These foundational themes remain influential, as evidenced by their central positioning and enduring connectivity (e.g., Bucholtz & Hall, 2005; Pessin et al., 2020). In contrast, the emergence of topics such as teacher identity, professional identity, and agency reflects a growing concern with lived experiences and professional practices, particularly in educational settings. This shift aligns with broader trends in applied linguistics that emphasize practitioner perspectives, reflexivity, and the situated nature of identity construction (Varghese et al., 2005; De Costa & Norton, 2017; Tsui, 2007).

The increasing visibility of English language teaching further suggests that it has become a key site for examining identity negotiation, likely due to globalization, mobility, and the expanding role of English in diverse sociocultural contexts (Murray, 2018). However, the relatively peripheral role of language policy, ideology, and power (considering the color and the size of nodes) indicates that structural and political dimensions of identity may remain underrepresented compared to classroom-based and professional identity research. Taken together, the findings suggest that while identity and language research is theoretically mature, future studies could benefit from reconnecting emerging, practice-oriented themes with broader ideological and policy-

oriented perspectives. Such integration would allow for more comprehensive accounts of how identities are shaped across individual, institutional, and sociopolitical levels.

### **The most influential documents**

The citation analysis of documents highlights the intellectual foundations and theoretical evolution of identity and language research. The prominence of Norton (1995) confirms its status as a landmark study that fundamentally reshaped the field by introducing the concept of investment as an alternative to traditional motivation frameworks in second language acquisition. The high citation count and strong linkage within the citation network suggest that this work continues to serve as a theoretical cornerstone for subsequent identity-related research. In addition, the strong influence of Heller (2003) and Bucholtz (1999) reflects the field's deep roots in sociolinguistic and poststructuralist traditions, which conceptualize identity as fluid, socially negotiated, and discursively constructed. These studies expanded the analytical lens of identity research by incorporating power relations, social positioning, and cultural practices into language-related identity formation. Furthermore, the inclusion of Morita (2004) and Phinney et al. (2001) demonstrates the field's expansion toward understanding identity in educational and intercultural contexts, particularly among second language learners and minority groups. These works suggest a growing recognition of identity as a multidimensional construct shaped by both linguistic experiences and broader sociocultural processes. In addition, the presence of Kanno and Stuart (2011) among the most influential works indicates an increasing scholarly focus on teacher identity and professional identity development. This

shift mirrors broader developments in applied linguistics that emphasize teacher agency, reflective practice, and identity negotiation within institutional and pedagogical settings.

Several influential studies, including Bucholtz (1999), Morita (2004), and Huffaker (2005), demonstrate high citation counts but limited direct citation links within the network. This pattern may indicate that these works exert influence across multiple disciplinary domains rather than being embedded within a tightly interconnected citation cluster. Such a distribution suggests the interdisciplinary nature of identity and language scholarship, which draws theoretical insights from fields such as sociology, psychology, communication studies, and education.

### **Theoretical and conceptual foundations underpin research on *identity and language***

The co-citation analysis of references reveals that identity and language research is theoretically grounded in four interrelated traditions: sociolinguistic and ideological perspectives on identity, learner identity and investment in language learning, dialogic and post-structural identity theories, and language teacher professional identity. The presence of a distinct dialogic cluster underscores the foundational role of dialogism in shaping how identity is conceptualized across the field. Dialogic theory, as articulated by Bakhtin (1981), provides an epistemological foundation for viewing identity as emergent, relational, and historically situated. Its strong co-citation ties with Norton's (2013) work and critical sociolinguistic scholarship suggest that dialogism enables researchers to bridge structure and agency, linking ideological forces with individual voice and participation.

The sociocultural and poststructuralist tradition, particularly Norton's (2013) identity framework, builds upon this dialogic

foundation by operationalizing identity in applied contexts such as language learning. Its central position in the network reflects its role as a theoretical mediator, translating abstract philosophical concepts into empirically tractable constructs such as investment, legitimacy, and community membership. Sociolinguistic theories continue to provide the macro-level grounding for identity research, emphasizing power, inequality, and ideological processes. Meanwhile, the teacher identity cluster illustrates how these foundational theories are extended into professional and institutional domains, where identity is negotiated within specific policy and pedagogical contexts.

Importantly, the yellow dialogic cluster's strong cross-cluster linkages indicate that identity and language research is characterized by theoretical convergence rather than compartmentalization. Dialogism operates as a shared conceptual resource that enables coherence across diverse research strands, supporting a unified yet pluralistic understanding of identity.

### **Conclusion and Implications**

This study set out to systematically map the development, intellectual structure, and emerging trends in research on language and identity through a comprehensive bibliometric analysis. By examining publication output, citation patterns, author productivity, thematic clustering, and co-citation networks, the study provides an overview of how identity and language scholarship have evolved over the past two decades. The findings demonstrate sustained growth in research output, reflecting the increasing centrality of identity as a construct in linguistic inquiry. At the same time, the analyses reveal a field characterized by theoretical diversity and thematic

expansion, with multiple research traditions contributing to its development.

The thematic analyses highlight several well-established research areas, including identity construction in language learning, multilingualism, and sociocultural and poststructuralist approaches. At the same time, emerging themes, particularly those related to teacher identity, suggest ongoing shifts in research priorities. Citation and co-citation analyses further reveal the foundational role of poststructuralist and sociocultural theories, while also indicating increasing cross-fertilization among applied linguistics, sociolinguistics, and related disciplines. However, the dispersion of influential works across theoretical traditions underscores the fragmented nature of the field.

The findings have several implications for future research on language and identity. First, the identified fragmentation highlights the need for greater theoretical integration across frameworks, particularly in bridging micro-level analyses of identity performance with macro-level examinations of ideology, power, and institutional structures. Second, emerging themes identified through overlay visualization suggest promising directions for future inquiry, including teacher identity, professional identity, and agency. Researchers are encouraged to engage with these areas using both qualitative and mixed-method approaches to strengthen cumulative knowledge building.

Methodologically, this study demonstrates the value of bibliometric approaches for synthesizing large and complex bodies of literature in language-related research. Practically, the findings can inform curriculum development, teacher education, and language policy by highlighting dominant ideologies and

emerging perspectives that shape understandings of identity in language use and learning.

## References

- Abourehab, Y., & Azaz, M. (2020). Pedagogical translanguaging in community/heritage Arabic language learning. *Journal of Multilingual and Multicultural Development*, 44(5), 398-411. <https://doi.org/10.1080/01434632.2020.1826496>
- Bakhtin, M. M. (1981). *The dialogic imagination: Four essays*. University of Texas Press.
- Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20(2), 107-128.
- Benzehaf, B. (2021). Multilingualism and its role in identity construction: A study of English students' perceptions. *International Journal of Multilingualism*, 20, 1145-1163. <https://doi.org/10.1080/14790718.2021.2003369>
- Berman, E., & Smith, B. (2021). De-naturalizing the novice: A critique of the theory of language socialization. *American Anthropologist*. <https://doi.org/10.1111/AMAN.13624>
- Block, D. (2007). *Second language identities*. Continuum.
- Block, D. (2013). The structure and agency dilemma in identity and intercultural communication research. *Language and Intercultural Communication*, 13, 126 - 147. <https://doi.org/10.1080/14708477.2013.770863>
- Blommaert, J. (2010). *A sociolinguistics of globalization*. Cambridge University Press
- Bose, P., Gao, X., Starfield, S., Sun, S., & Ramdani, J. M. (2023). Conceptualization of family and language practice in family language policy research on migrants: A systematic review. *Language Policy*, 22(3), 343-365. <https://doi.org/10.1007/s10993-023-09661-8>
- Bourdieu, P. (2001). Le mystère du ministère. *Actes de la recherche en sciences sociales*, 140(5), 7-11.
- Bucholtz, M. (1999). Why be normal: Language and identity practices in a community of nerd girls. *Language in Society*, 28(2), 203-223.
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4), 585-614.
- Darvin, R. & Norton, B. (2015). Identity and a model of investment in applied linguistics.

- Annual Review of Applied Linguistics, 35, 36-56.  
<https://doi.org/10.1017/S0267190514000191>
- De Costa, P. I., & Norton, B. (2017). Introduction: Identity, transdisciplinarity, and the good language teacher. *The Modern Language Journal*, 101(S1), 3-14.
- Dennison, J. E., Rusnalasari, Z. D., Marselina, S., & Rafli (2025). Language ideologies and policy development: Navigating identity, education, and globalization. *Lingua*, 3(1), 62-76.  
<https://doi.org/10.61978/lingua.v3i1.1038>
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research and Policy*, 10(1), 45-53.  
<https://doi.org/10.1177/135581960501000110>
- Duff, P. A., & Uchida, Y. (1997). The negotiation of teachers' sociocultural identities and practices in postsecondary EFL classrooms. *TESOL Quarterly*, 31(3), 451-486.
- Evnitskaya, N., & Morton, T. (2011). Knowledge construction, meaning-making, and interaction in CLIL science classroom communities of practice. *Language and Education*, 25(2), 109-127.  
<https://doi.org/10.1080/09500782.2010.547199>
- Fisher, L., Evans, M., Forbes, K., Gayton, A., & Liu, Y. (2018). Participative multilingual identity construction in the languages classroom: A multi-theoretical conceptualization. *International Journal of Multilingualism*, 17(4), 448-466.  
<https://doi.org/10.1080/14790718.2018.1524896>
- Giles, H., & Johnson, P. (1987). Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language*, (68), 69-99.
- Han, M., H. (2025). Korean transnational poetry and the negotiation of identity: Korean, English, and Konglish. *International Journal of Bilingualism*, 13670069251363822.  
<https://doi.org/10.1177/13670069251363822>
- Heller, M. (2003). Globalization, the new economy, and the commodification of language and identity. *Journal of Sociolinguistics*, 7(4), 473-492.
- Henry, A (2017). L2 motivation and multilingual identities. *The Modern Language Journal*, 101, 548-565.  
<https://doi.org/10.1111/MODL.12412>
- Hofer, B. (2025). The multicompetent mind - how does early multilingualism affect children's language development and cognition? *Intercultural Education*, 36(5), 548-563.  
<https://doi.org/10.1080/14675986.2025.2514398>
- Humphrey, M., Laycock, D. & Umbach M. (2019). Introduction. *Journal of Political Ideologies*, 24, 113 - 120.  
<https://doi.org/10.1080/13569317.2019.1589961>
- Irvine, J. T., & Gal, S. (2000). Language ideology and linguistic differentiation. In P. V. Kroskrity (Ed.), *Regimes of language: Ideologies, politics, and identities* (pp. 35-84). School of American Research Press..
- Jamali, D., Samara, G., Lokman, I. M. (2023). Determinants of research productivity and efficiency among the Arab world's accredited business schools. *Management Review Quarterly*, 74, 2511 - 2543.  
<https://doi.org/10.1007/s11301-023-00365-1>
- Joseph, J. E. (2004). *Language in national identities*. Palgrave Macmillan UK.
- Kanno, Y. (2003). Imagined communities, school visions, and the education of bilingual students in Japan. *Journal of Language, Identity, and Education*, 2(4), 285-300.
- Kanno, Y., & Stuart, C. (2011). Learning to become a second language teacher: Identities-in-practice. *Modern Language Journal*, 95(2), 236-252.
- Kaplan, A, Garner, J (2017). A complex dynamic systems perspective on identity and its development: The dynamic systems model of role identity. *Developmental Psychology*, 53, 2036-2051.  
<https://doi.org/10.1037/dev0000339>
- Kinginger, C. (2011). Language, identity, and study abroad: Sociocultural perspectives. *International Journal of Bilingual Education and Bilingualism*, 14, 625 - 627.  
<https://doi.org/10.1080/13670050.2010.538263>
- Kiramba, L., K. (2018). Language ideologies and epistemic exclusion. *Language and Education*, 32, 291-312.  
<https://doi.org/10.1080/09500782.2018.1438469>
- Larsen-Freeman, D. (2018). Looking ahead: Future directions in, and future research into, second language acquisition. *Foreign Language*

- Annals, 51(1), 55-72.  
<https://doi.org/10.1111/flan.12314>
- Li, A. & Roberts, G. (2023). Co-occurrence, extension, and social salience: The emergence of indexicality in an artificial language. *Cognitive science*, 47(5), e13290.  
<https://doi.org/10.1111/cogs.13290>
- Li, M. (2025). Identity in second language learning: Theoretical perspectives and research trends with an emphasis on Chinese as a second language. *Journal of Literature and Language*, 1(1), 50-65.  
<https://doi.org/10.71204/jhxdxz72>
- Liu, Q., & Chao, C.-C. (2017). CALL from an ecological perspective: How a teacher perceives affordance and fosters learner agency in a technology-mediated language classroom. *ReCALL*, 30(1), 68-87.  
<https://doi.org/10.1017/s0958344017000222>
- Liu, Y. & Yueting, X. (2011). Inclusion or exclusion? a narrative inquiry of a language teacher's identity experience in the 'new work order' of competing pedagogies. *Teaching and Teacher Education*, 27(3), 589-597.
- Mazzucchelli, T., Burton, E., & Roberts, L. (2019). Scholarly productivity and citation impact of Australian academic psychologists. *Australian Journal of Psychology*, 71, 305-311.  
<https://doi.org/10.1111/AJPY.12248>
- McKeown, S., & Mir, Z. M. (2021). Considerations for conducting systematic reviews: Evaluating the performance of different methods for de-duplicating references. *Systematic Reviews*, 10(1).  
<https://doi.org/10.1186/s13643-021-01583-y>
- Moradi, H. (2020). An exploration of autonomy, agency, and identity in EFL contexts. Springer.
- Mori, M. (2014). Conflicting ideologies and language policy in adult ESL: Complexities of language socialization in a majority-L1 classroom. *Journal of Language, Identity and Education*, 13, 153-170.  
<https://doi.org/10.1080/15348458.2014.919810>
- Morita, N. (2004). Negotiating participation and identity in second language academic communities. *TESOL Quarterly*, 38(4), 573-603.
- Murray, D. (2018). The world of English language teaching: Creating equity or inequity? *Language Teaching Research*, 24, 60 - 70.  
<https://doi.org/10.1177/1362168818777529>
- Nie, S. (2025). Language hybridity as psychological resistance: Identity construction in native American literature from a social psychology perspective. *Environment and Social Psychology*, 10(10), 1-24.  
<https://doi.org/10.59429/csp.v10i10.4117>
- Norton, B. (1995). Social identity, investment, and language learning. *TESOL Quarterly*, 29(1), 9-31.
- Norton, B. (1997). Language, identity, and the ownership of English. *TESOL Quarterly*, 31(3), 409-429.
- Norton, B. (2000). Identity and language learning: Gender, ethnicity and educational change. Longman.
- Norton, B. (2013). Identity and language learning: Extending the conversation. *Multilingual matters*.
- Norton, B. (2015). Identity, investment, and faces of English internationally. *Chinese Journal of Applied Linguistics*, 38, 375 - 391.  
<https://doi.org/10.1515/CJAL-2015-0025>
- Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics*, 28(2), 163-188.
- Pavlenko, A., & Blackledge, A. (2004). Negotiation of Identities in multilingual contexts. *Multilingual Matters*.
- Perlman, A. J. (2020). Southeast Asian refugee-learners: Identities informing ESL education and support. *Journal of Southeast Asian American Education and Advancement*, 15(1).  
<https://doi.org/10.7771/2153-8999.1179>
- Pessin, V. Z., Yamane, L. H., & Siman, R. R. (2022). Smart bibliometrics: An integrated method of science mapping and bibliometric analysis. *Scientometrics*, 127(6), 3695-3718.  
<https://doi.org/10.1007/s11192-022-04406-6>
- Phinney, J., Romero, I., Nava, M., & Huang, D. (2001). The role of language, parents, and peers in ethnic identity among adolescents in immigrant families. *Journal of Youth and Adolescence*, 30(2), 135-153.
- Rozenvalde, K., & Martena, S. (2024). The value of language skills in Latvia: From socio-economic inequality to multilingual awareness. *Sociolinguistica*, 38(1), 1-26.  
<https://doi.org/10.1515/soci-2022-0037>
- Sah, P., K. (2024). Teachers' beliefs and reproduction of language ideologies in English-medium instruction programs in Nepal. *International Journal of Bilingualism*, 28, 701-718.  
<https://doi.org/10.1177/13670069241236701>
- Schiller, N. G. (2012). Situating identities: Towards identity studies without binaries of difference. *Identities*, 19, 520-532.

- <https://doi.org/10.1080/1070289X.2012.741525>
- Sultana, S. (2022). Translingual practices and national identity mediated in the semiotized digital spaces. *Australian Review of Applied Linguistics*, 45(3). <https://doi.org/10.1075/aral.21051.sul>
- Sung, C. C. M. (2019). Negotiating participation and identity in a second language: Mainland Chinese students' English learning experiences in a multilingual university in Hong Kong. *Research Papers in Education*, 36(4), 381-401. <https://doi.org/10.1080/02671522.2019.1677760>
- Tong, Y., Hong, Y. & Chiu, S., C. (1999). Language use as a carrier of social identity. *International Journal of Intercultural Relations*, 23, 281-296. [https://doi.org/10.1016/S0147-1767\(98\)00039-X](https://doi.org/10.1016/S0147-1767(98)00039-X)
- Tsui, A. B. (2007). Complexities of identity formation: A narrative inquiry of an EFL teacher. *TESOL Quarterly*, 41(4), 657-680.
- Ushioda, E. (2017). The impact of global English on motivation to learn other languages: Toward an ideal multilingual self. *The Modern Language Journal*, 101(3), 469-482. Portico. <https://doi.org/10.1111/modl.12413>
- Varghese, M., Morgan, B., Johnston, B., & Johnson, K. A. (2005). Theorizing language teacher identity: Three perspectives and beyond. *Journal of Language, Identity, and Education*, 4(1), 21-44.
- Vélez-Rendón, G. (2010). From social identity to professional identity: Issues of language and gender. *Foreign Language Annals*, 43, 635-649. <https://doi.org/10.1111/j.1944-9720.2010.01113.x>
- Wei, L. (2017). Translanguaging as a practical theory of language. *Applied Linguistics*, 39, 9-30. <https://doi.org/10.1093/APPLIN/AMX039>
- Wenger, E. (2009). *Communities of practice: The key to knowledge strategy*. Routledge.
- Yung, K. W. H. (2025). Theorizing agency from a complex dynamic systems theory perspective: Evidence from language learner narratives in online shadow education. *The Modern Language Journal*, 109(1), 274-294. Portico. <https://doi.org/10.1111/modl.12988>
- Zembylas, M. (2003). Emotions and teacher identity: A poststructural perspective. *Teachers and Teaching*, 9, 213-238. <https://doi.org/10.1080/13540600309378>
- Zhang, L., & Glänzel, W. (2012). Proceedings papers in journals versus the regular journal publications. *Journal of Informetrics*, 6, 88-96. <https://doi.org/10.1016/j.joi.2011.06.007>