ISSN: 2770-4602

JLT Journal of Language Teaching

AN INTERNATIONAL JOURNAL OF
LANGUAGE TEACHING AND RESEARCH

Vol. 3, Issue 3
2023

JLT Journal of Language Teaching

ISSN (online): 2770-4602
Vol. 3, Issue 3, 2023

CONTENTS

Review

How Educational Intercultural Bilingualism has shaped indigenous people language learning in Mexico: Decolonial challenges **1-7**
Xushan Wei

A CiteSpace-based review of cross-cultural communication in international Chinese language education in China (1994-2022) **8-18**
Chunyan Ma, Shuang Zhang, Ruixin Wang, Ting Yi, Chili Li

A critical understanding of second language acquisition from two sociolinguistic strands: The Variationist approach and the Investment Perspective **19-25**
Chang Liu

Original Research

A corpus-based study to evaluate the generativist explanation of children's error patterns in questions **26-38**
Yiran Du

Please visit our website to read more articles for free: https://jlt.ac/

© Pioneer Publications LTD 2021-2023

Journal of Language Teaching, *3*(3), 1-7, 2023
https://doi.org/10.54475/jlt.2023.003

Check for updates

Review

How Educational Intercultural Bilingualism has shaped indigenous people language learning in Mexico: Decolonial challenges

Xushan Wei

University of Durham, Durham, United Kingdom

Received: January 15, 2023 / Accepted: March 1, 2023 / Published Online: March 3, 2023
© Pioneer Publications LTD 2023

Abstract

Educational Intercultural Bilingualism (EIB), this approach to language education in multi-ethnic integration countries has received attention from a wide range of experts and scholars. This is because the right to be educated in one's mother tongue and to speak a dialect based on indigenous people often does not coexist with the official language of cultural identity at the national level. Better inclusion and valuing the place of indigenous people in the national cultural identity makes language education a social movement for equity and adaptation progress.

This paper will focus on the EIB in Mexico, through a multi-level analysis of the historical processes and challenges encountered in language education for indigenous populations decolonization. The richness of Mexico's linguistic diversity and changing social relations will be useful in helping to build on-board experience in language teaching around the world.

Keywords language learning; Mexico; Educational Intercultural Bilingualism (EIB); indigenous people; decolonization

1. Introduction

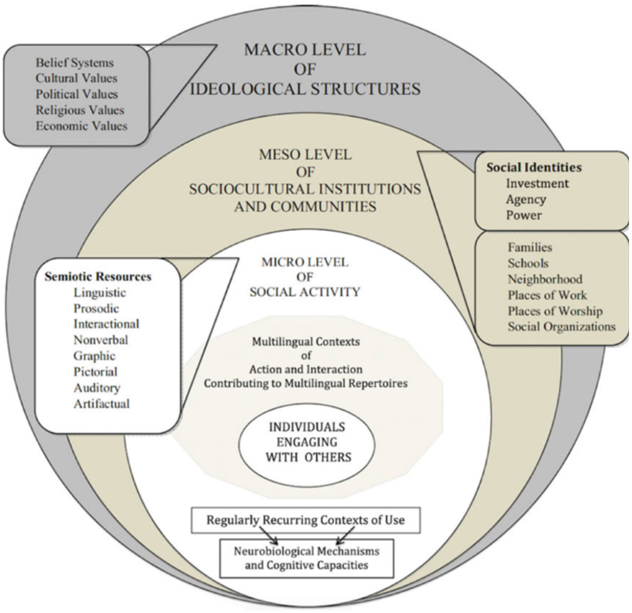
After more than three hundred years of Spanish colonisation, Mexico has been confronted with the linguistic obstacles of decolonization, prompting broad issues in intercultural language instruction. This report aims to expand on Duff (2019)'s model of "The Multifaceted Nature of Language Learning and Teaching" by examining the three dimensions of macro, meso, and micro in relation to Blommaert (2010)'s theoretical framework on language and mobility in examine the changing forms of "local" and "translocal" social identities. The focus is on the impact of Educational Intercultural Bilingualism (EIB) (López, 2021) on language learning among indigenous people in Mexico who have survived for a lengthy moment in such a society where Spanish is the official language and are subjected to socially inequitable top-down policies. In the process of learning and utilising Spanish, they have had to conceal their indigenous identity (Aman, 2017; Guerrettaz, 2020; Messing, 2007; O'Donnell, 2010). The language use of indigenous populations is marked by 'Diglossia' (Ferguson, 1959), which signifies that bilinguals who speak both languages prefer to switch between situations. As a result of the Mexican government reclaiming the right to decolonization, schools at every level became crucial spaces for altering social order and identity. This study explores how the EIB influenced language acquisition for indigenous groups and assisted them in the process while they were under pressure. It has revealed the shifting policies of indigenous groups, exploring the structural characteristics that surroundings support and influence, and placing the causes of their profound suffering into perspective. The analysis is followed by a realistic relevance to the difficulties and a forecast of the evolution of intercultural language education within the context of Covid-19 (New Coronary Pneumonia).

2. Theories and context language(s) background

2.1. Duff's model: process of language learning and teaching

As depicted in Figure 1, this model analyses the social concepts and influences on the language learning process, particularly second language acquisition (SLA), at various levels. Key points are made regarding the need for learners to have equal opportunities for language learning and to communicate their cultures in a diverse environment, avoiding the effects of the world's worrisome political systems and discursive restrictions, as well as providing guidelines for language learning research on indigenous decolonization. Avoiding discrimination and prejudice in racial exclusion and focusing on recognising linguistic variety would be preferable. This research uses the model's three aspects to a study of language learning and instruction among indigenous groups in Mexico.

Figure 1. The Multifaceted Nature of Language Learning and Teaching (Duff, 2019)


Belief Systems
Cultural Values
Political Values
Religious Values
Economic Values
MACRO LEVEL OF IDEOLOGICAL STRUCTURES
MESO LEVEL OF SOCIOCULTURAL INSTITUTIONS AND COMMUNITIES
Social Identities
Investment
Agency
Power
Families
Schools
Neighborhood
Places of Work
Places of Worship
Social Organizations
Semiotic Resources
Linguistic
Prosodic
Interactional
Nonverbal
Graphic
Pictorial
Auditory
Artifactual
MICRO LEVEL OF SOCIAL ACTIVITY
Multilingual Contexts of Action and Interaction Contributing to Multilingual Repertoires
INDIVIDUALS ENGAGING WITH OTHERS
Regularly Recurring Contexts of Use
Neurobiological Mechanisms and Cognitive Capacities

2.2. Sociolinguistic scales

With the process of the modernization and globalisation, indigenous people of Mexico have steadily migrated from independent dwelling regions to metropolitan areas. Faced with changes in the system of mobilitation, the 'local' and 'translocal' ideas in Blommaert (2010)'s 'Sociolinguistic scale' (Table 1) determine the transitions that occur between migrant populations. This explains the condition of indigenous peoples, whose use of the official language and adaptation to local customs constitute the 'local' shift, whereas their 'translocal' habituation under the policy of intercultural bilingualism in 'code-switching' maintains their capacity to communicate with their families in the indigenous language and to carry out social activities in Spanish. From the dimensions of time and space, the social relationship between the two languages is reflected in the distribution of power-relationships.

Table 1. The general direction of people's move (Blommaert, 2010)

	Lower scale	Higher scale
Time	Momentary	Timeless
Space	Local, situated	Translocal, widespread

2.3. 'Diglossia'

Ferguson (1959) compared four countries with multilingualism in the middle of the twentieth century (Table 2) by comparing the official language and the dialect (native language) of the country, and concluded that countries with colonies generally divide the use of language into different scenarios. The classification of 'diglossia' is as follows: High variety denoted by 'H' for official or majority language and Low variety denoted by 'L' for indigenous language (regional dialect). In European countries, this style of linguistic expression is collectively referred to as "bilingual." Table 3 displays the patterns of indigenous 'diglossia' language use in Mexico based on 'diglossia'.

Table 2. Examples of diglossia in 4 countries (Ferguson, 1959)

		H is called	L is called
Arabic	Classical (=H)	*'al-fuṣḥā*	*'al-ʿāmmiyyah, 'ad-dārij*
	Egyptian (=L)	*'il-faṣīḥ, 'in-naḥawi*	*'il-ʿammiyya*
Sw. German	Stand. German (=H)	*Schriftsprache*	*Hoochtüütsch*
	Swiss (=L)	*[Schweizer] Dialekt, Schweizerdeutsch*	*Schwyzertüütsch*
H. Creole	French (=H)	*français*	*créole haïtien*
Greek	H and L	*katharévusa*	*dhimotikí*

Table 3. Indigenous people language use in Mexico

Context	Indigenous population (American Indians) in Mexico
Language background	Diglossia (Ferguson, 1959) Spanish (H)+ indigenous language (L) 'H' official language use in public 'L' indigenous language (regional dialect) use in informal areas

3. Background of context

The linguistic community of Mexico is comprised of a variety of indigenous languages and the official language. In a web-based data survey reported by Vega (2022), in 'Instituto Nacional de Estadística, Geografía, e Informática' (INEGI) Agency 2010, it is shown that Mexico still maintains a status of 68 indigenous languages and more than three hundred derived languages and dialects of 'Sociolinguistic scales' of indigenous groups (Blommaert, 2010), and he addresses seven of the more widespread indigenous languages. In Table 4, I compared the number of language speakers to the total population of the country, from which it can be inferred that a certain percentage of the indigenous population are proficient speakers of indigenous languages, even though Náhuatl is considered by Vega (2022) to be the second official language of Mexico due to its widespread circulation. In contrast to the inferior status of indigenous languages, Spanish is viewed as the dominant language of the coloniser and serves as a proof of identity within the society.

Table 4. The most spoken indigenous language in Mexico

Rankings	Indigenous language	Number of speakers	Total population	Percentage
1	Náhuatl	1,500,000	114,000,000	1.32%
2	Mayan (Yucatecan)	780,000	114,000,000	0.68%
3	Mixteco	477,000	114,000,000	0.42%
4	Zapoteco	450,000	114,000,000	0.39%
5	Tzeltal	445,000	114,000,000	0.39%
6	Tzoltzil	400,000	114,000,000	0.35%
7	Otomí	280,000	114,000,000	0.25%

4. Main text

Indigenous populations reside primarily in the hills, deserts, and coastlines, even so, with modernisation and globalisation, they have begun to migrate towards agricultural, industrial, and urban areas. The authority of the official language had formed a phenomenon that weakened indigenous groups' sense of culture and language by forcing their integration into higher education and assimilation of languages (López-Gopar et al., 2021; O'Donnell, 2010; Tinajero & Englander, 2011). It is demonstrated that indigenous communities are disadvantaged and that their educational environment is complicated and changeable. The discussion that follows is organised around Duff (2019)'s model of language acquisition and education, analysing how did EIB shaped the issues of decolonising the languages of indigenous populations at the stage of political, social, and individual level.

4.1. Macro level: societal values — Colonial power and alliances (Duff, 2019)

From a macro perspective, the colonial-relation changes in the implementation of EIB in Latin America are reflected in López (2021)'s study, which proves that language could gradually achieve assimilationist ideas in the process of cultural education, thereby adjusting the inherent development of colonial relations at the social level.

Chronologically, the more than three hundred years of Spanish colonisation of Mexico, beginning in the 16th century, left the indigenous people marginalised in terms of quality of life and language, resulting in an absence of superior learning resources for indigenous students in a state of social discrimination. In 1821, when Mexico attained independence and founded a nation, decolonization began. Under liberal reforms, indigenous people were granted the ability to be legally 'equal citizens' in 1857 (Mendoza Zuany, 2009). Prior to the Mexican Revolution of 1921, the national government seriously ignored the education of indigenous children (Tinajero & Englander, 2011). Since the 20th century, governments in Latin America have implemented the EIB strategy to assist minimize ethnic differences among indigenous populations, with the goal of encouraging more indigenous people to acquire the official language and fostering the assimilation of language use in society. This procedure, however, has neglected the necessity for indigenous people to utilise and conserve their own languages, as well as the need to protect indigenous languages through averting cultural loss. The fears of the citizens of Spanish monolingualism that indigenous access to education would unseat the privileges and status of citizens, as well as the opposition expressed in the early years of the development of EIB academic achievement (1940s), exemplify the conflicting social relations that have always existed in the process of language education. To prevent such inequities and foster harmonious societal norms, Mexico organised government-level consultative forums (López, 2021). In the 1940s, the Mexican government has set up the 'Indigenous Institute' with the intention of integrating indigenous populations into the acculturation of the Mexican nation. Under the government's top-down policy, Spanish-only primary schools emerged, which compelled students and teachers to use Spanish and prohibited the use of indigenous languages, preventing the transmission of indigenous languages to the next generation and causing problems of exclusion for indigenous groups. In these schools where the official language was taught, indigenous youngsters suffered from the poor teaching quality provided by city-based teachers and the cultural bias of the curriculum(Mendoza Zuany, 2009). During the 1960s, the Mexican colonial language policy was still in the exploratory phase of language use(de Suarez, 1973), and there were numerous issues and flaws during this time. The approach influenced the use of indigenous languages as a means of learning official languages and the acknowledgement of bilingualism for indigenous students on a national basis (Hamel & Francis, 2006). The legal application of EIB by the Mexican government provided a swift push for the gradual transformation of Spanishization's homogeneous approach into bilingualism and interculturality. The State clearly acknowledged Mexico as a multicultural and multilingual nation in its laws, and EIB programmes were widely implemented in schools of all stages to foster the dual identity of indigenous learners as national and equal citizens (Paciotto, 2004; Tinajero & Englander, 2011). O'Donnell (2010) reports that indigenous students consistently have a substantially lower participation rate in higher education than urban populations, and that their academic achievement in school is lower than that of other students. During the Spanish colonisation of Chihuahua in the sixteenth century, the indigenous language of the Tarahu-mara people was severely threatened, and their ethnic group's land area was reduced to half of its original size(Paciotto, 2004). This questionable oppression has resulted in the marginalisation of indigenous languages, and its effects are still recognized presently. In 2011, the government reformed the constitution on indigenous language settings, preserving and developing the normative permit for indigenous language rights in schools and society (López, 2021), affirming the state's obligation to guarantee and expand EIB in the expectation of integrating cultural differences and reducing discrimination against minority indigenous peoples.

In the recent era, EIB schooling has helped indigenous people to gradually move out of the countryside into modern life with the changes of language mastery and habitual adaptations. Investigating the language use of indigenous populations and the effects of state political discourse and educational policies has received attention (Tinajero & Englander, 2011), as has the emergence of indigenous people becoming monolingual speakers of predominantly official languages and gradually adopting assimilation. In this setting, indigenous communities in Mexico have perpetually unequal social relationships, were materially impoverished, had limited access to higher education, and continued to be discriminated against (López-Gopar et al., 2021).

4.2. Meso level: the role of schools, institutions and communities (Duff, 2019)

From a meso level standpoint, Mexico lost the majority of its indigenous languages during European colonisation, however today's language revitalisation is exemplified by schools, which have launched a global movement to decolonize languages. In an effort to prevent their loss, schools have become more conscious of the need to help revive indigenous languages in recent years. As we encounter the twenty-first century, the Mexican government has enacted a constitution for universalist education that supports the concurrent study of Spanish and indigenous languages in schools; conversely, the formation of indigenous bilingual education models is limited by the lack of teacher capacity. Poverty is the primary cause of this situation (Tinajero & Englander, 2011). Some studies have evaluated the current state of multilingualism in Mexico, anticipating the gradual loss of indigenous languages in the lack of educational support in indigenous communities and even predicting that they could become extinct in this century (O'Donnell, 2010).

In curriculum planning for Mexico's primary schools, the education of indigenous children up to the age of 12 is a specific instance that deviates from the policies of the regular system of education. This stage is the first opportunity for indigenous children to be exposed to normative language teaching away from their original home education environment. The school curriculum, the recruitment of teachers, and educational policies form a separate system, and institutions in remote areas are even less developed. The process of trial and error is fraught with the search for an educational model that is suitable for indigenous students through a variety of approaches (Tinajero & Englander, 2011). The Bilingual Bicultural Education Program (BBEP) policy was implemented in the Chihuahua region in 1991, and background research for Paciotto (2004)'s study of indigenous students and bilingual education in rural primary schools revealed that language use was locally differentiated according to function, for home environments and community settings. The majority of interactions occur in the indigenous language, but Spanish is preferred in school and social situations. Parents and teachers consider school as the first opportunity for indigenous children to be exposed to a diverse culture and learn the capacity to speak and write Spanish. In this circumstance, the learning policy is committed to maintaining and fostering the circulation of indigenous languages and respecting the existence of indigenous cultures. In terms of facilitating the integration of younger children into a bilingual environment, guidance is conducted exclusively in the mother tongue (indigenous language) in the first year, with oral training in Spanish added after the second year, followed by a gradual shift of the main classroom activities to Spanish after the third year, with materials designed to be as bilingual as possible. There is still opposition to this type of EIB programme, with some students having high dropout and low enrollment rates as a result of negative family attitudes and financial constraints. Hamel and Francis (2006) conducted a study on bilingualism in indigenous primary schools in the state of Michoacan, using indigenous languages as a model for the excellent integration of Spanish as a second language into the classroom. The study reveals the need to address the opening of higher education opportunities at the national level at the middle and high school levels as an incentive for graduates to learn, once the classroom model has been established, in order to assist indigenous groups in retaining their native languages while being willing to learn Spanish.

The issue of intercultural ethnicity in secondary education in Mexico for indigenous middle and high school students is crucial to political concerns and institutional inclinations. In 1993, the inclusion of secondary education in basic education was explicitly included in Mexican law, and the educational transition of public schools for indigenous students was put in place, with the alternative integration of cultural elements of identity in the curriculum, helping to strengthen the identity of indigenous youth and enabling the dialogue of cultural difference to serve out through actual cultural activities.

At the higher education level, the 1960s student movements in Mexico prompted a geopolitical reflection on decoloniality (Fúnez-Flores, 2022), the student perspective being that of a group that perceived the educational system's weak points, and the resistance movements that arose accelerating social processes and providing an ideal for restructuring. The formation of resistance movements accelerates social processes and generates optimal responses to the concept of globalising the social structure. In recent years, the combined efforts of government and educational institutions have produced encouraging results: a study by Mendoza Zuany (2009), which interviewed teachers and students from different regions, noted that teachers need to spend more time instructing indigenous students, based on the nature of intercultural universities, in relation to the process of crossing original cultures and languages. The study recommends that schools acquire an understanding of indigenous knowledge in order to more effectively address the practical challenges that arise with their students. O'Donnell (2010) conducted a comparative study of two groups of monolingual Spanish-speaking students and bilingual indigenous language learners in bilingual Mexican universities. The study demonstrates that university students who are bilingual in indigenous languages and Spanish have an advantage in acquiring skills in English as a third language, indicating that language acquisition is not the only disadvantage for indigenous people in the expansion of globalisation, but that there is a positive impact in terms of greater access to social opportunities and excellent adaptation skills that can help to challenge inequalities. López-Gopar et al. (2021) also investigated the process of learning English for Mexican learners by assisting three students from indigenous backgrounds with low socioeconomic status (SES). They suggest that multilingual classroom approaches should be utilised to escape the crippling inequities rooted in monolingual homogeneity. In response to the Aboriginal university training approach, Dietz (2009) proposes extending the bilingual education model to the elementary level.

4.3. Micro level: individual identity changes & social activities (Duff, 2019)

Mexico had 92 languages and 62 distinct indigenous communities as of the 1990 census (Tinajero & Englander, 2011). According to a census conducted in 2008, there were 68 indigenous groups, around 12 million Mexicans could be classified as indigenous based on their language use and ethnicity, and sixty percent of the indigenous population spoke only indigenous languages. However, indigenous languages are rapidly fading owing to indigenous communities' lack of educational support. Individual problems at the micro level include lack of access to quality education and social disparities between rich and poor, with private bilingual schools and English-learning higher education institutions in Mexico excluding more than 95% of the population and the persistence of racist colonialism in Mexican society (López-Gopar et al., 2021) English-learning-based higher education is contributing to the construction of social justice and promoting multilingualism while confronting class concerns, gender issues, and colonial inequities. The UNESCO Guidelines for Intercultural Education of 2006 provide an overview of multicultural and intercultural education (Unesco Education, 2006). Inside of it, it is suggested, in relation to the minority cultures in which indigenous people belong, that indigenous peoples are defined as groups living in specific social, cultural, and economic conditions, often requiring government regulation of language and customs in accordance with distinct political institutions, and that 'indigenous' identity is both self-identified and identified by others. This guidance suggested that education systems apply effective and appropriate programmes for indigenous groups to promote the acquisition of knowledge and skills that will help them in better incorporating into the culture and economy of their societies, paying special attention to indigenous women and children as well as migrants. Messing (2007) examined the ideology of identity views of indigenous communities in Mexico and discovered a state of denigration of indigenous identities and misunderstanding of the transitions between tradition and modernity among indigenous populations. In the transformed of language use, it was proposed that the impact of social shift on the self-concept of indigenous populations should always be considered. The process of language transition and extinction in the Tlaxcala region has highlighted that code switching is not only an individual process, but also a societal and institutional change. Aman (2017) believes that based on the findings of the culture of the Indian Highlands, that at the geopolitical level, colonised indigenous communities were required to speak the colonial language in an official bilingual learning environment. Indigenous groups would view the usage of their language in bilingual classrooms as a stigma. Consequently, power relations determine the development of cultural diversity. Guerrettaz (2020) investigated the identity of students in the Yucatec region of Mexico towards their original race, the 'Maya,' and concluded that Mexicans suffer from a persistent identity crisis and wish to restore or preserve their ancestral tongue. As a consequence of their investigation, it was determined that indigenous language revival practitioners have not actually been merged with the objective of addressing postcolonialism. After entering higher education, indigenous students are forced to relinquish more of their indigenous culture, conceal their ethnic identity, and change their official language in order to advance further up the educational ladder and be treated on a par with their monolingual Mexican-speaking urban peers (O'Donnell, 2010). Nieto (2018) addresses the challenge of decolonization by analysing the discourse on civic education in several Latin American nations, highlighting the positive role of multilateral institutions in constructing the desired goals of educational reform, and proposing that collective consciousness cannot be ignored and that unjust and structural global configurations of rights can be avoided epistemologically.

During the Covid-19 epidemic, school systems in a number of nations faced challenges. The Mexican model of intercultural education had both beneficial and negative consequences on online education, and Dietz and Cortés (2021) explain the changes for indigenous children in Veracruz after the closing of bilingual schools. In March 2020, following the closure of schools as a result of the epidemic, the government began exploring the development of an online teaching model. However, due to the wide variety of indigenous languages, the majority of online course content chose Spanish as the only official curriculum language, forcing indigenous students to return to monolingual Spanish. To prevent the spread of the virus, this is the framework of a top-down approach that disregards the bilingual learning environment of indigenous pupils. Therefore, students from poorer areas are unable to continue taking online classes, and others are compelled to drop out of school for economic reasons, widening the gap between native students and Spanish speakers. As a result of the establishment of Covid-19, new shortcomings of top-down policies have been found, as well as the need to further design and strengthen the flexibility of language teaching in the classroom and change the school's long-term management model following the return to normalcy.

5. Conclusion and recommendation

The findings of this research provide insights to examine the influences of EIB in language use of indigenous peoples in Mexico by reviewing the history of the EIB's development and research on the process of policy implementation, the role of institutions and indigenous groups. Drawing on Duff (2019)'s model of language education, this study evaluates the EIB's influence in different dimensions using literature data. The study finds that, in response to top-down government policies, teaching institutions at all levels have used structural features to play a supportive role in teaching official languages and preserving indigenous languages, while also innovating curriculum design at various stages of history to advance the recognition of language decolonisation. Teaching and learning endangered indigenous languages requires additional policy support and methodological research from a power perspective. The scope of this study was limited in terms of raw data from the context of indigenous people, it would be benefit forward with more empirical investigation to enhance the progression and would be a fruitful are for further research on language education.

Xushan Wei is an MA in Education graduate from the School of Education, Durham University, United Kingdom. Her research interests include the study of pedagogical concepts and the study of international language education. She has a particular interest in educational issues in the international context. She has worked on a study of Kazakhstan students' language learning and intercultural adaptation during their study in China, and the articles had been published in. various academic journals.
Email: weixushan2020@163.com

References

Aman, R. (2017). Colonial differences in intercultural education: On interculturality in the Andes and the decolonization of intercultural dialogue. *Comparative Education Review*, *61*(S1), S103-S120.

Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press.

de Suarez, Y. L. (1973). Shirley Brice Heath, Telling tongues: Language policy in Mexico, colony to nation. New York and London: Teachers College Press, 1972. Pp. xx + 300. *Language in Society*, *2*(2), 294-297. https://doi.org/10.1017/S0047404500000774

Dietz, G. (2009). Intercultural universities in Mexico: empowering indigenous peoples or mainstreaming multiculturalism? *Intercultural Education*, *20*(1), 1-4. https://doi.org/10.1080/14675980802700623

Dietz, G., & Cortés, L. S. M. (2021). Mexican intercultural education in times of COVID-19 pandemic. *Intercultural Education*, *32*(1), 100-107. https://doi.org/10.1080/14675986.2020.1843895

Duff, P. A. (2019). Social dimensions and processes in second language acquisition: Multilingual socialization in transnational contexts. *The Modern Language Journal*, *103*, 6-22.

Ferguson, C. A. (1959). Diglossia. *word*, *15*(2), 325-340.

Fúnez-Flores, J. I. (2022). Toward decolonial globalisation studies. *Globalisation, Societies and Education*, 1-21. https://doi.org/10.1080/14767724.2022.2048796

Guerrettaz, A. M. (2020). “We are the mayas”: Indigenous language revitalization, identification, and postcolonialism in the Yucatan, Mexico. *Linguistics and Education*, *58*, 100765. https://doi.org/10.1016/j.linged.2019.100765

Hamel, R. E., & Francis, N. (2006). The Teaching of Spanish as a Second Language in an Indigenous Bilingual Intercultural Curriculum. *Language, Culture and Curriculum*, *19*(2), 171-188. https://doi.org/10.1080/07908310608668761

López, L. E. (2021). What is educación intercultural bilingüe in Latin America nowadays: results and challenges. *Journal of Multilingual and Multicultural Development*, *42*(10), 955-968. https://doi.org/10.1080/01434632.2020.1827646

López-Gopar, M. E., Schissel, J. L., Leung, C., & Morales, J. (2021). Co-constructing Social Justice: Language Educators Challenging Colonial Practices in Mexico. *Applied Linguistics*, *42*(6), 1097-1109. https://doi.org/10.1093/applin/amab047

Mendoza Zuany, R. G. (2009). Building hybrid knowledge at the Intercultural University of Veracruz, Mexico: an anthropological study of indigenous contexts. *Intercultural Education*, *20*(3), 211-218. https://doi.org/10.1080/14675980903138558

Messing, J. (2007). Multiple ideologies and competing discourses: Language shift in Tlaxcala, Mexico. *Language in Society*, *36*(4), 555-577. https://doi.org/10.1017/S0047404507070443

Nieto, D. (2018). Citizenship education discourses in Latin America: multilateral institutions and the decolonial challenge. *Compare: A Journal of Comparative and International Education*, *48*(3), 432-450. https://doi.org/10.1080/03057925.2017.1408399

O'Donnell, J. L. (2010). The indigenous, national, and international language in higher education: students' academic trajectories in Oaxaca, Mexico. *International Journal of Applied Linguistics*, *20*(3), 386-416. https://doi.org/10.1111/j.1473-4192.2010.00254.x

Paciotto, C. (2004). Language Policy, Indigenous Languages and the Village School: A Study of Bilingual Education for the Tarahumara of Northern Mexico. *International Journal of Bilingual Education and Bilingualism*, *7*(6), 529-548. https://doi.org/10.1080/13670050408667829

Tinajero, G., & Englander, K. (2011). Bilingual-intercultural education for indigenous children: the case of Mexico in an era of globalization and uprisings. *Intercultural Education*, *22*(3), 163-178. https://doi.org/10.1080/14675986.2011.592019

Unesco Education, S. (2006). *UNESCO guidelines on intercultural education*. UNESCO.

Vega, E. (2022). Map of the most spoken indigenous languages in Mexico. *[online] Xcaret Blog*. https://doi.org/https://blog.xcaret.com/en/map-of-the-most-spoken-indigenous-languages-in-mexico/

Journal of Language Teaching, *3*(3), 8-18, 2023
https://doi.org/10.54475/jlt.2023.004

Review

A CiteSpace-based review of cross-cultural communication in international Chinese language education in China (1994-2022)

Chunyan Ma[1], Shuang Zhang[2], Ting Yi[3], Ruixin Wang[1], Chili Li[4]

1. New Era University College, Kajang, Selangor, Malaysia
2. School of Humanities, Wuhan University of Engineering Science, Wuhan, China
3. Research Institute of International Chinese Language Education, Beijing Language and Culture University, Beijing, China
4. School of Foreign Languages, Hubei University of Technology, Wuhan, China

Received: January 26, 2023/ Accepted: February 19, 2023 / Published Online: March 1, 2023
© Pioneer Publications LTD 2023

Abstract

This paper applies bibliometrics visualization analysis software CiteSpace-6.1.6 and CNKI as the data source and analyzes the literature published in domestic journals during 1994-2022 in China, in order to understand the development trend and main issues of cross-cultural communication in the field of international Chinese language education. The results of bibliometric cluster analysis show that in the past 30 years, the research topics in this field mainly focus on 17 clusters, including "*culture*", "*Chinese language*", "*cultural communication*", "*differences*", "*international students*", "*cultural differences*", "*Chinese as a foreign language*", "*cultural connotation*", "*foreign language teaching*" and "*cross-cultural*". According to the analysis of cited information, "*cultural difference*", "*pragmatic mistakes*", "*Chinese as a foreign language*", "*cultural communication*", "*cross-culture*", "*cultural teaching*" and "*cultural difference*" are the most concerned topics. "*Case study*" and "*teaching design*" are the research trends and hot spots in this field in the past two years. This paper not only reviews the history of cross-cultural studies in international Chinese language education, but also looks forward to the future development of cross-cultural studies in international Chinese language education.

Keywords international Chinese language education; cross-cultural communication; emerging trends; bibliometrics; visual analysis

1. Introduction

In recent years, the research on international Chinese language education has been fruitful, and the research theories and methods are constantly evolving (Li and Zhai, 2021). With Chinese language learned by more and more people (Ma et al., 2022; Gong et al., 2020), there are increasing cross-cultural communication problems in international Chinese language education. The number of Chinese as an additional language learners has exceeded 20 million by the end of 2020 (Li et al., 2021; Li et al., 2022). The International Conference on Chinese Language Education was held in Changsha in 2019, and the China Foundation for International Chinese Language Education was established in 2020. From the original "teaching Chinese as a foreign language" to the current "international Chinese language education", although international Chinese language education is a relatively new term, it can be used to refer to both the cause of international Chinese language education and the discipline. It is of great significance because it involves both the national language governance ability and the international influence of Chinese (Wang, 2021a). The name change of international Chinese language education brings a broader perspective to the discussion of teaching research. In the communication and interaction with different cultures around the world, we should think and discuss how to deal with the relationship between ourselves and others in the cross-cultural context (Wang, 2021b). Zhao (2014) proposed "Cross-cultural communication is generally regarded as a teaching method and a learning strategy". In the 21st century, the cultivation of intercultural communicative competence is the teaching goal of international Chinese language education, which is the requirement of social development for language talents and the inevitable trend of the development of international Chinese language education (Cui, 2022a;

Cui, 2022b; Zu, 2017; Zhang, 2017; Wang, 2015; Li et al., 2022). Therefore, cross-cultural studies play an important role in international Chinese language education. The goal of the curriculum for teaching Chinese as a foreign language in *the General Curriculum for Teaching Chinese as a Foreign Language* (2019) is the comprehensive ability to use Chinese language. The strategy and cultural ability in the four major contents (language skills, language knowledge, strategy and cultural ability) both mention cross-cultural communication. In view of this, this paper collects and collates all relevant studies on international Chinese language education research from the cross-cultural perspective of CNKI.

Compared with previous studies, this study has the following extensions and contributions:

1. The extension of research content: It gives statistical analysis of cross-cultural studies in international Chinese language education from the aspects of the number of literature, the distribution of literature sources, core authors and hot topics.
2. It is the expansion of the data cycle: the literature of nearly 30 years is collected to analyze the current situation of cross-cultural research in international Chinese language education in a longer time span.
3. The expansion of the database: with CNKI as the data source, the database is currently the largest full-text database in China.
4. The expansion of research methods: Visual mapping knowledge and bibliometric methods in CiteSpace 6.1.6 software are used to visually display the research hotspot, latest trend and evolution characteristics of cross-cultural communication.

2. Research Design

2.1. Research Questions

This paper addresses the following two research questions:

1. In the past 30 years, what is the overall development trend of international Chinese language education from the perspective of cross-cultural communication?
2. In the past 30 years, what are the research hotspots and main findings in the field of international Chinese language education from cross-cultural perspectives?

2.2. Data Source

The source of literature is from China National Knowledge Internet (CNKI). Literature collection and analysis includes three steps: retrieval, screening and analysis. Combined with the research theme, this paper defines the theme as the interdisciplinary research of international Chinese language education and intercultural communication. The search subject words in CNKI were "*Chinese language and Intercultural communication*", "*Teaching Chinese as a Foreign Language and Cross-cultural Communication*", "*International Chinese Language Education and Intercultural Communication*", and the time range is from the earliest to December 31, 2022. By setting the previous conditions, a total of 1850 literature records were retrieved. The conference papers were manually removed and finally sorted into 1,807 documents that met the conditions. The derived data were in Refwork format. There is no relevant literature before 1994, so the time span of the documents was from January 1, 1994 to December 31, 2022.

2.3. Analysis Tool

Literature visualization analysis software CiteSpace-6.1.6 was used in this study. At present, there are many kinds of document metrology software, such as VOSviewer, CiteSpace, Bibioshiny, HiteCite etc. By reading the literature comparing several software programs (Fu and Ding, 2019; Song and Chi, 2016; Chu and Zhang, 2019; Liao, 2011; Zhang et al., 2011, Hou and Hu, 2013), I tried several programs and finally decided to use CiteSpace. This software has certain advantages in data algorithm and clustering algorithm more in line with the requirements of this paper, more flexible data language (Both Chinese and English can be analyzed), and visual presentation methods of time and graph, especially when revealing the discipline research regular and research direction (Fu and Ding, 2019).

As a new method and a new field of scientometrics, scientific knowledge map or knowledge map is emerging and developing rapidly in academia. The science knowledge map is targeted at knowledge domain, and it is a kind of image that shows the relationship between the development process and structure of science knowledge with Knowledge domain as the object. It has the dual nature and characteristics of 'graph' and 'spectrum': it is not only a visual knowledge graph, but also a serialized knowledge lineage, showing the network, structure, interaction, cross, evolution or derivation among knowledge units or knowledge groups, and these complex knowledge relations are breeding the generation of new knowledge" (Chen et al., 2015; Chen, 2006).

This paper firstly makes descriptive statistics and analysis of cross-cultural studies in the field of international Chinese language education, calculates the time distribution of the number of journal publications, and uses CiteSpace 6.1.6 to conduct visual analysis of relevant studies. In the section of literature analysis, firstly, the bibliometric analysis software Citespace is used to handle basic data, and then the literature is summarized and explored in depth by clustering. Secondly, key journals of outstanding publication volume, active scholars and academic institutions were presented by statistical ranking of keyword frequency. Thirdly, the key words co-occurrence function of Citespace was used to execute cluster analysis on the core topics according to the size and distribution of nodes in the co-occurrence graph (Van Eck and Waltman, 2017). Finally, combined with the atlas data, the original literature is further studied, and the overall development trend of the research is comprehensively sorted out and discussed.

3. Results and Discussion

3.1. Annual Publication Amount

Firstly, the annual statistics of literature information are made. In the past 30 years, there have been more than 1,800 cross-cultural related papers in international Chinese language education, and the annual average publication amount is about 60. The annual publication amount is shown in Figure 1.

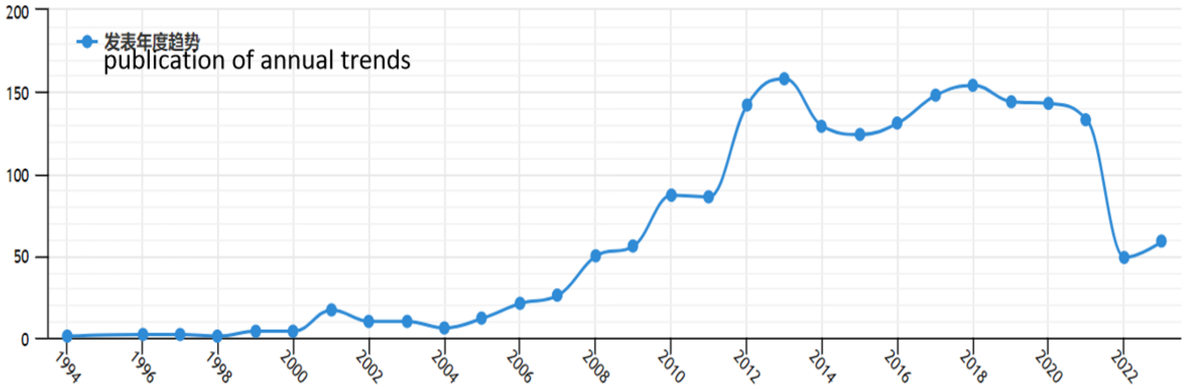


Figure 1. Annual Publication Amount (1994-2022)

As can be seen from Figure 1, the period from 1994 to 2007 was in the embryonic stage, and the number of relevant studies was relatively small, but the trend was rising slowly. The period from 2008 to 2021 was a booming period, with the number of relevant research surging and starting to decline in 2021. However, there was a new development trend in 2022, and the research trend began to rise again. The number of publications reached a peak of more than 160 in 2013.

In 2008, due to the improvement of China's comprehensive strength, the global Chinese language learning boom began. With the increasing number of Chinese learning groups, cross-cultural studies in international Chinese language education have become increasingly rich and developed rapidly with the development of "Chinese fever". However, the number of studies dropped rapidly in 2021, which could be closely related to the COVID-19 outbreak in 2019 and China's epidemic prevention policies. Due to the epidemic, most Chinese learners can only learn online instead of studying in China, which to some extent avoids some communication problems. As the increasing international migration of the 21st century has generated enormous social and cultural implications in the host countries (Peng et al., 2021), and cultural differences caused by cultural conflicts. Therefore, the number of papers published in 2022 dropped to about 50.

3.2. Authors

CiteSpace software was used to make statistics on the authors of literatures. Among the 1807 literatures examined, there were 533 authors (including the first author and co-authors), and three authors who published 3 relevant papers. 48 of them published two papers. The statistics of all authors of 3 papers and some authors of 2 papers are shown in Table 1 according to the publication frequency.

Table 1. Publication Frequency of Authors

No.	Author	Freq.	No.	Author	Freq.
1	Xiao Zhikui	3	10	Tan Ruwei	2
2	Wu Leya	3	11	Wang Yongyang	2
3	Liu Wei	3	12	Liu Li	2
4	Lin Tian	2	13	Liu Siyang	2
5	Guo Guangwei	2	14	Jiang Lili	2
6	Liu Wei	2	15	Zhang Xiaohong	2
7	Chen Lei	2	16	Gu Yutong	2
8	Wang Peng	2	17	Cui Youwei	2
9	Zhang Ning	2	18	Li Hong	2

CiteSpace software was used to process the authors of the literature, and the visualization map of co-occurrence of authors was obtained, as shown in Figure 2.

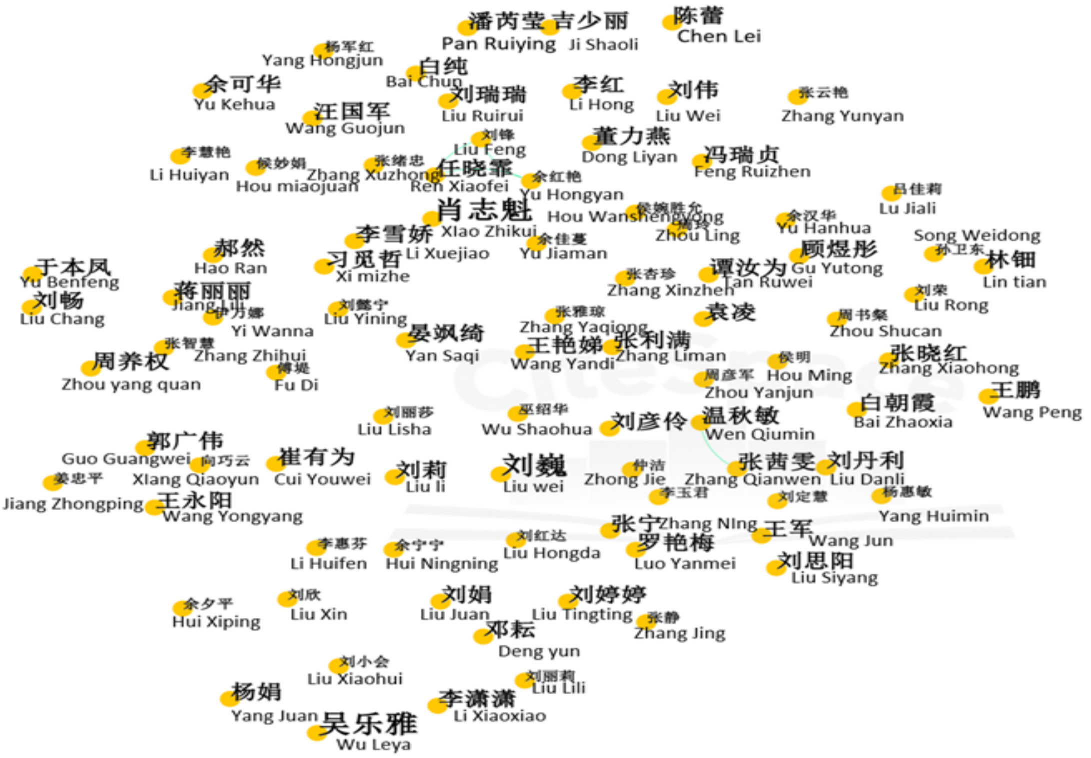

潘芮莹 Pan Ruiying
吉少丽 Ji Shaoli
陈蕾 Chen Lei
杨军红 Yang Hongjun
白纯 Bai Chun
余可华 Yu Kehua
刘瑞瑞 Liu Ruirui
李红 Li Hong
刘伟 Liu Wei
张云艳 Zhang Yunyan
汪国军 Wang Guojun
刘锋 Liu Feng
董力燕 Dong Liyan
冯瑞贞 Feng Ruizhen
李慧艳 Li Huiyan
候妙娟 Hou miaojuan
张绪忠 Zhang Xuzhong
任晓霏 Ren Xiaofei
余红艳 Yu Hongyan
吕佳莉 Lu Jiali
肖志魁 Xlao Zhikui
侯婉胜允 Hou Wanshengyong
周玲 Zhou Ling
余汉华 Yu Hanhua
Song Weidong
李雪娇 Li Xuejiao
余佳蔓 Yu Jiaman
顾煜彤 Gu Yutong
孙卫东
林钿 Lin tian
于本凤 Yu Benfeng
郝然 Hao Ran
习觅哲 Xi mizhe
张杏珍 Zhang Xinzhen
谭汝为 Tan Ruwei
刘荣 Liu Rong
刘畅 Liu Chang
蒋丽丽 Jiang Lili
伊万娜 Yi Wanna
刘懿宁 Liu Yining
张雅琼 Zhang Yaqiong
袁凌
周书粲 Zhou Shucan
晏飒绮 Yan Saqi
王艳娣 Wang Yandi
张利满 Zhang Liman
张晓红 Zhang Xiaohong
周养权 Zhou yang quan
张智慧 Zhang Zhihui
傅堤 Fu Di
侯明 Hou Ming
周彦军 Zhou Yanjun
王鹏 Wang Peng
刘丽莎 Liu Lisha
巫绍华 Wu Shaohua
刘彦伶
温秋敏 Wen Qiumin
白朝霞 Bai Zhaoxia
郭广伟 Guo Guangwei
向巧云 Xlang Qiaoyun
崔有为 Cui Youwei
刘莉 Liu li
刘巍 Liu wei
仲洁 Zhong Jie
张茜雯 Zhang Qianwen
刘丹利 Liu Danli
姜忠平 Jiang Zhongping
王永阳 Wang Yongyang
李玉君
刘定慧
杨惠敏 Yang Huimin
张宁 Zhang NIng
王军 Wang Jun
李惠芬 Li Huifen
余宁宁 Hui Ningning
刘红达 Liu Hongda
罗艳梅 Luo Yanmei
刘思阳 Liu Siyang
余夕平 Hui Xiping
刘欣 Liu Xin
刘娟 Liu Juan
刘婷婷 Liu Tingting
张静 Zhang Jing
邓耘 Deng yun
刘小会 Liu Xiaohui
刘丽莉 Liu Lili
杨娟 Yang Juan
李潇潇 Li Xiaoxiao
吴乐雅 Wu Leya

Figure 2. Visual Map of Co-occurrence by Authors

As can be seen from Table 1, among the literature analyzed, Xiao Zhikui from Shanghai International Studies University, Wu Leya from Communication University of China, and Liu Wei from Beijing Youth University of Political Science published the most papers, each with 3 papers. At the same time, the co-occurrence map of authors in Figure 2 shows that although there are some co-published papers among a few authors, in general, researchers present a scattered distribution and a long-term stable academic community has not been formed. This not only shows that in the international Chinese language education, as a young science, cross-cultural research is a greatly promising field. The characteristics of the early development stage of "fighting alone" and "spreading in many areas" also mean that it is urgent to strengthen the construction of specialized academic organizations, academic conferences, academic journals and other disciplines in this field. It is necessary to actively promote the policy reform of the current academic evaluation mechanism and encourage interdisciplinary, cross-school and cross-industry cooperation. Moreover, the relationship between international Chinese language education and cross-cultural education is inseparable.

3.3. Author Affiliations

CiteSpace software was used to visualize the literature data sources, and the co-occurrence map of the institutions to which the authors are affiliated was obtained in Figure 3, as shown below.

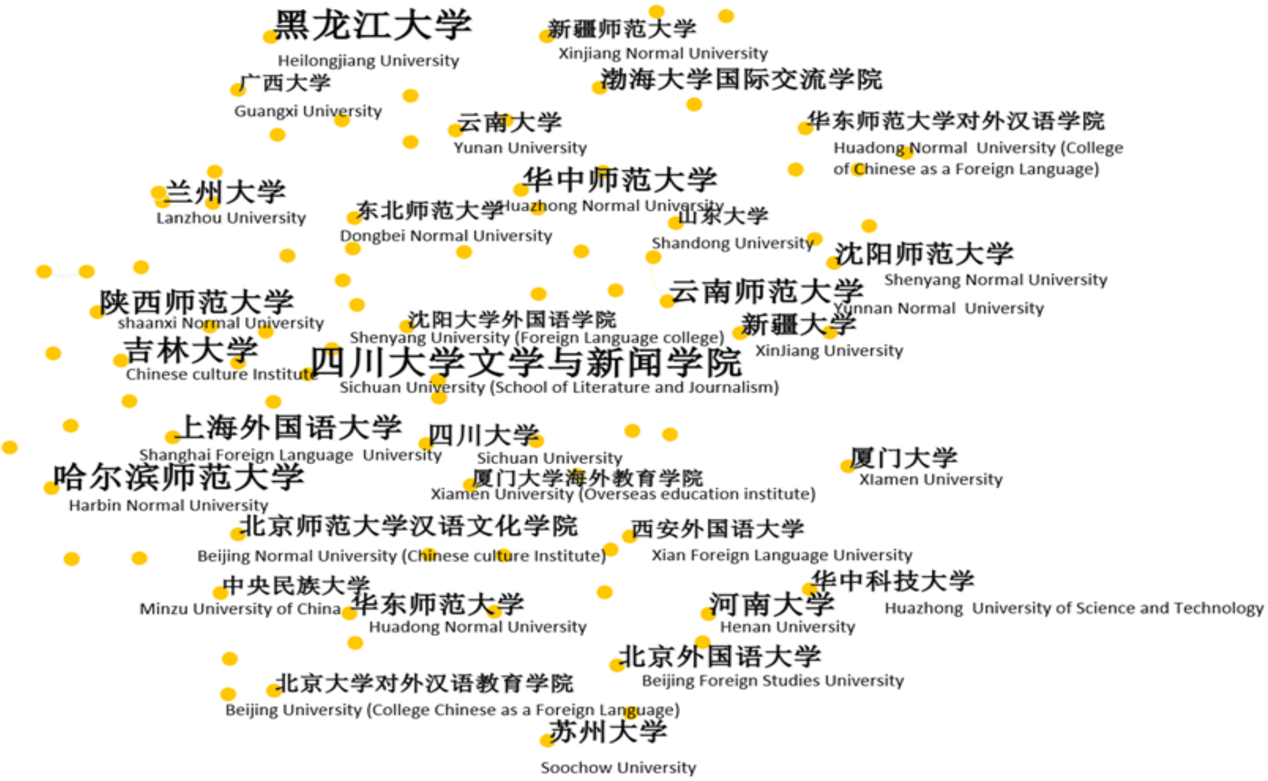

黑龙江大学
Heilongjiang University
新疆师范大学
Xinjiang Normal University
渤海大学国际交流学院
广西大学
Guangxi University
云南大学
Yunan University
华东师范大学对外汉语学院
Huadong Normal University (College of Chinese as a Foreign Language)
华中师范大学
Huazhong Normal University
兰州大学
Lanzhou University
东北师范大学
Dongbei Normal University
山东大学
Shandong University
沈阳师范大学
Shenyang Normal University
云南师范大学
Yunnan Normal University
陕西师范大学
shaanxi Normal University
沈阳大学外国语学院
Shenyang University (Foreign Language college)
新疆大学
XinJiang University
吉林大学
Chinese culture Institute
四川大学文学与新闻学院
Sichuan University (School of Literature and Journalism)
上海外国语大学
Shanghai Foreign Language University
四川大学
Sichuan University
厦门大学
Xiamen University
哈尔滨师范大学
Harbin Normal University
厦门大学海外教育学院
Xiamen University (Overseas education institute)
北京师范大学汉语文化学院
Beijing Normal University (Chinese culture Institute)
西安外国语大学
Xian Foreign Language University
中央民族大学
Minzu University of China
华东师范大学
Huadong Normal University
华中科技大学
Huazhong University of Science and Technology
河南大学
Henan University
北京外国语大学
Beijing Foreign Studies University
北京大学对外汉语教育学院
Beijing University (College Chinese as a Foreign Language)
苏州大学
Soochow University

Figure 3.Visual Map of Author Affiliations

As can be seen from Figure 3, most of the literature data comes from universities, such as Heilongjiang University, Xinjiang Normal University, Soochow University, Sichuan University, Peking University, Jilin University and Beijing Foreign Studies University. A small number of the literature data comes from non-university middle schools and technical secondary schools, such as Wuzhou Teachers' College, Xianning Hot Spring Middle School and Tunxi No. 1 Middle School.

3.4. Research Hotspots

Keywords in literature extract and summarize research objects and core ideas. Geng and Gao (2022) concluded that "the occurrence frequency and co-occurrence of keywords can reveal research hotspots and central topics in a certain field". We imported the exported literature data into CiteSpace 6.1.6 software and used CiteSpace to conduct a preliminary analysis of keyword frequency and aggregation of all literature. Table 2 shows the keywords with 9 or more frequency, and Figure 3 shows the keyword clustering map, listing the main 12 keyword clustering. Table 2 and Figure 4 show the main research hotspots of intercultural communication in the field of international Chinese language education.

Table 2. List of High-frequency Keywords

No.	High-frequency keywords	Freq.	N	High-frequency keywords	Freq.
1	Culture difference	102	19	Case study	18
2	Pragmatic mistakes	98	20	Euphemism	17
3	Chinese as a foreign language	94	21	Teaching	17
4	Culture communication	66	22	Foreign language teaching	16
5	Culture	62	23	Communication barrier	16
6	Cross-culture	56	24	Cultural connotation	16
7	Difference	34	25	Chinglish	16
8	Translation	32	26	Contrast	15
9	Culture teaching	31	27	Communication	15
10	Strategy	29	28	Translation strategy	13
11	Language	29	29	Pragmatics	12

12	International student	28	30	Countermeasure	12
13	Cultural conflict	25	31	Chinese	11
14	Chinese	25	32	Chinese culture	11
15	English teaching	22	33	Teaching strategy	11
16	Chinese teaching	22	34	Cultivation	11
17	English	21	35	Communicative competence	10
18	Politeness principle	19	36	Culture lead-in	9

Visualization processing of keywords is carried out to get Figure 4, keyword co-occurrence map. Keywords and the relationship between keywords are displayed in an intuitive visual way, as shown below.

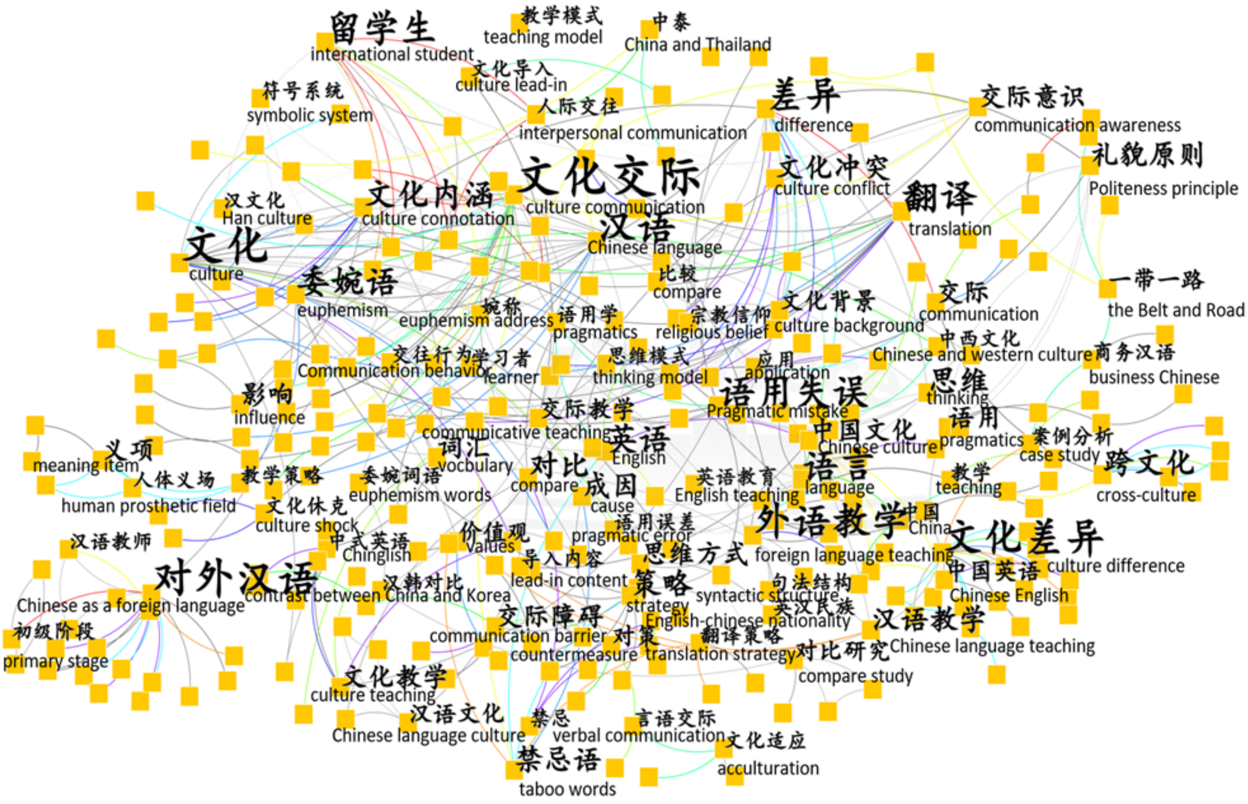

留学生
international student
教学模式
teaching model
中泰
China and Thailand
文化导入
culture lead-in
符号系统
symbolic system
人际交往
interpersonal communication
差异
difference
交际意识
communication awareness
礼貌原则
Politeness principle
文化冲突
culture conflict
翻译
translation
汉文化
Han culture
文化内涵
culture connotation
文化交际
culture communication
汉语
Chinese language
文化
culture
委婉语
euphemism
比較
compare
婉称
euphemism address
语用学
pragmatics
宗教信仰
religious belief
文化背景
culture background
交际
communication
一带一路
the Belt and Road
中西文化
Chinese and western culture
商务汉语
business Chinese
交往行为
Communication behavior
学习者
learner
思维模式
thinking model
应用
application
语用失误
Pragmatic mistake
思维
thinking
影响
influence
交际教学
communicative teaching
英语
English
中国文化
Chinese culture
语用
pragmatics
案例分析
case study
跨文化
cross-culture
义项
meaning item
人体义场
human prosthetic field
教学策略
委婉词语
euphemism words
词汇
vocbulary
对比
compare
成因
cause
英语教育
English teaching
语言
language
教学
teaching
文化休克
culture shock
语用误差
pragmatic error
外语教学
foreign language teaching
中国
China
文化差异
culture difference
汉语教师
中式英语
Chinglish
价值观
Values
导入内容
lead-in content
思维方式
中国英语
Chinese English
对外汉语
Chinese as a foreign language
汉韩对比
contrast between China and Korea
策略
strategy
句法结构
syntactic structure
英汉民族
English-chinese nationality
汉语教学
Chinese language teaching
初级阶段
primary stage
交际障碍
communication barrier
对策
countermeasure
翻译策略
translation strategy
对比研究
compare study
文化教学
culture teaching
汉语文化
Chinese language culture
禁忌
言语交际
verbal communication
文化适应
acculturation
禁忌语
taboo words

Figure 4. Visualization of Keywords Clustering

As can be seen from Table 2 and Figure 4, "cultural difference" is the keyword with the highest frequency, appearing 102 times, while "pragmatic mistakes" and "Chinese as a foreign language", two topics closely related to international Chinese language education, are next with the frequency of 98 and 94 times respectively. The most prominent feature of international Chinese language education is always conducting the Chinese teaching under different cultural backgrounds, in which many key words are related to culture, such as "culture", "cultural differences", "cultural communication", "cross-culture", "cultural conflict", "cultural connotation", "Chinese culture" and "cultural introduction". The statistical data more directly reflect the importance of intercultural communication in international Chinese language education, as well as the close relationship between intercultural communication and Chinese teaching. Some other high-frequency keywords are also basically around the three main keywords, from the three aspects of cultural differences, pragmatics and Chinese teaching, in-depth discussion of cross-cultural issues in international Chinese language education from a more micro perspective, such as translation errors, euphemism expression, politeness principle, teaching strategies and communicative strategies. This study did not combine synonyms in the statistical analysis. The purpose of this study is not only to reflect the slightly different research perspectives of experts in the academic circle, but also to illustrate the current situation that consensus and unified standards have not been formed on the use of some basic concepts and terms in the academic field.

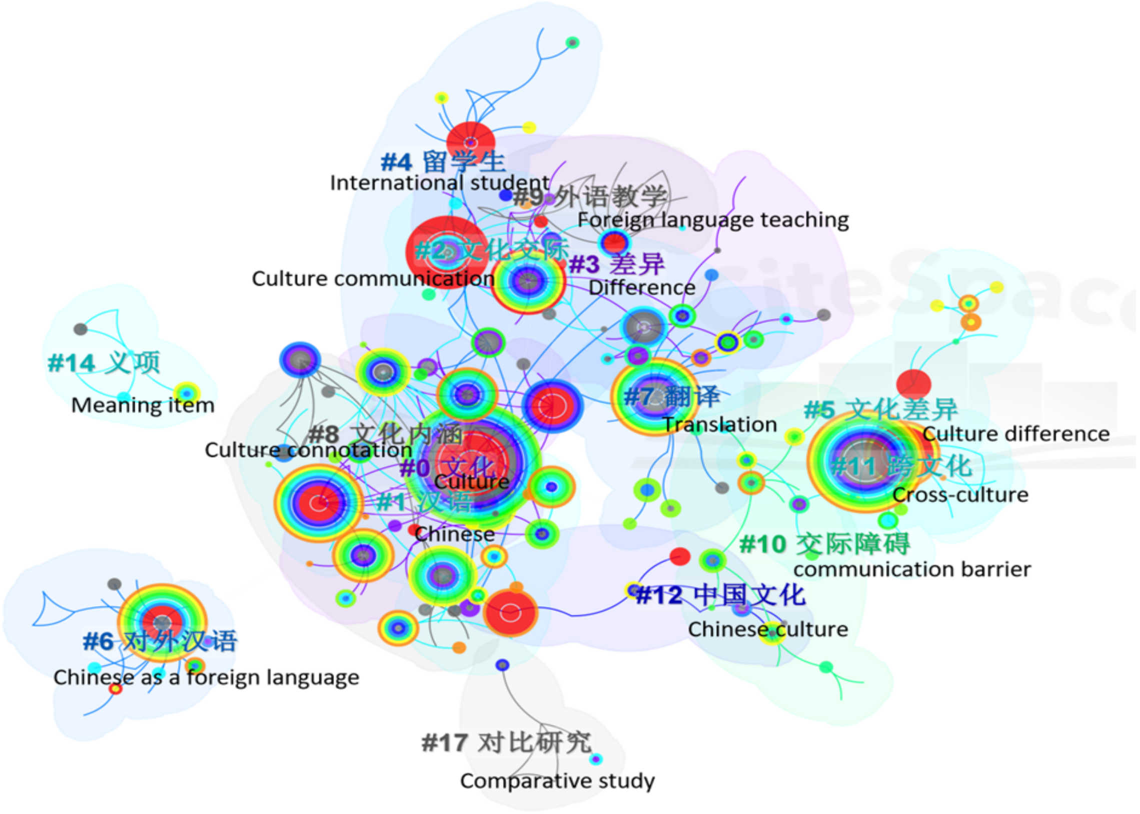

#4 留学生
International student
#9 外语教学
Foreign language teaching
#2 文化交际
Culture communication
#3 差异
Difference
#14 义项
Meaning item
#8 文化内涵
Culture connotation
#0 文化
Culture
#1 汉语
Chinese
#7 翻译
Translation
#5 文化差异
Culture difference
#11 跨文化
Cross-culture
#10 交际障碍
communication barrier
#12 中国文化
Chinese culture
#6 对外汉语
Chinese as a foreign language
#17 对比研究
Comparative study

Figure 5. Clustering Map of Main Keywords

Through the keyword clustering processing, as shown in Figure 5, 12 keyword clusters are shown, such as "culture", "Chinese language", "cultural communication", "differences", "international students", "cultural differences", "Chinese as a foreign language", "cultural connotation", "foreign language teaching" and "cross-culture". It can be seen from the keyword cluster map that "culture"and "Chinese language" are the two most important hot spots, which also proves once again that the two sciences of international Chinese language education and intercultural communication are permeated and connected.

3.5. Research Trends

The keyword clustering shows the overall research hotspot between 1994 and 2022. In order to further explore the internal changes in the popularity of relevant research topics in the past 30 years, we use the time-line clustering analysis function based on terms provided by CiteSpace to investigate the text information more deeply and comprehensively. Geng and Gao (2022) mentioned that this function can carry out text processing on the title and abstract keywords of the literature and form a term base by using TF-IDF (word frequency-inverse document frequency) algorithm, and then carry out co-occurrence correlation analysis on the terms of different articles, so as to more comprehensively reveal the main cluster of the study and the specific research dynamic changes of the cluster at each time node (in this study, it is every year).

In addition, "CiteSpace software also provides the 'burst term' analysis function based on terms, through which the preface dynamics and development trend of a certain research field can be analyzed. In other words, words with high frequency change rate and fast frequency growth rate are detected in a specified time period (See Figure 6). Different from the time line analysis function, emergent words reflect the relative change amplitude rather than the absolute number of word frequency of a certain term, so they can be used to observe the preface dynamics and development trends of research topics (Geng and Gao, 2022). CiteSpace 6.1.6 is used to analyze the emergence of cross-cultural communication in the field of international Chinese language education from 1994 to 2022. Figure 7 shows the time line development of 17 clusters and the distribution of the most significant 12 burst words.

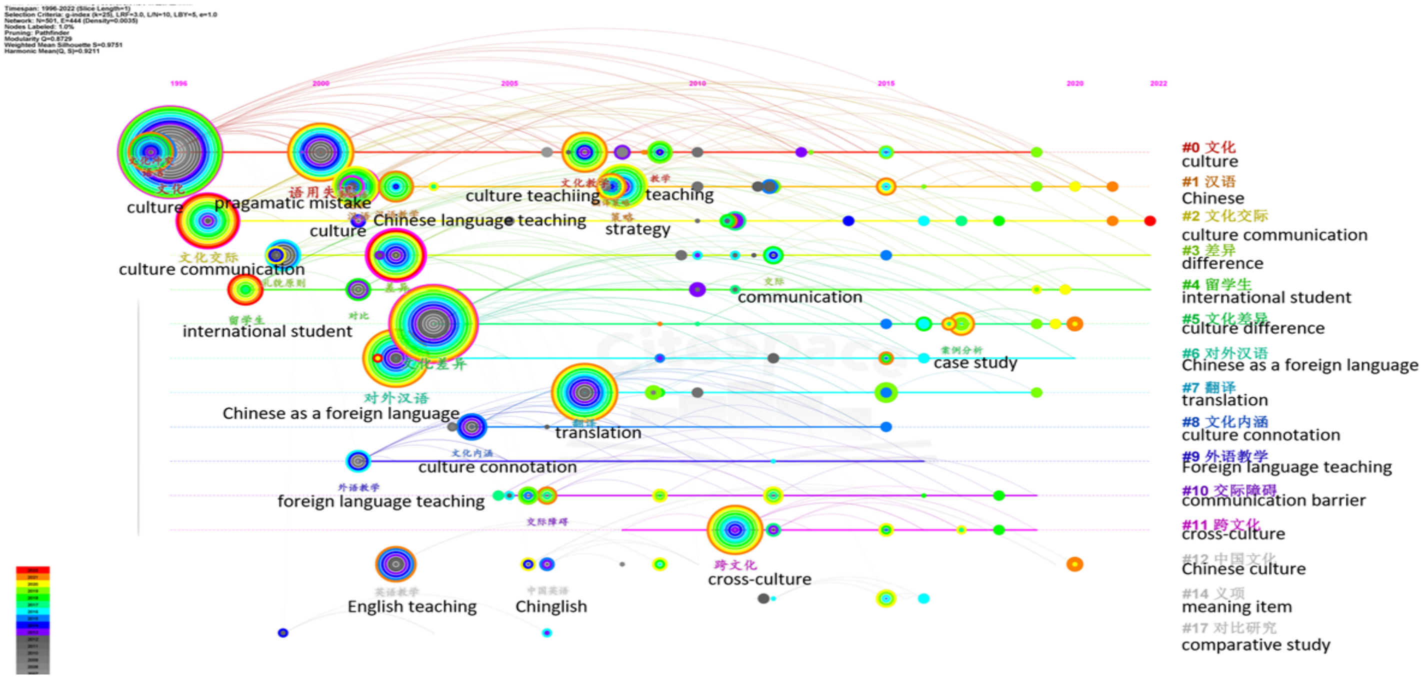
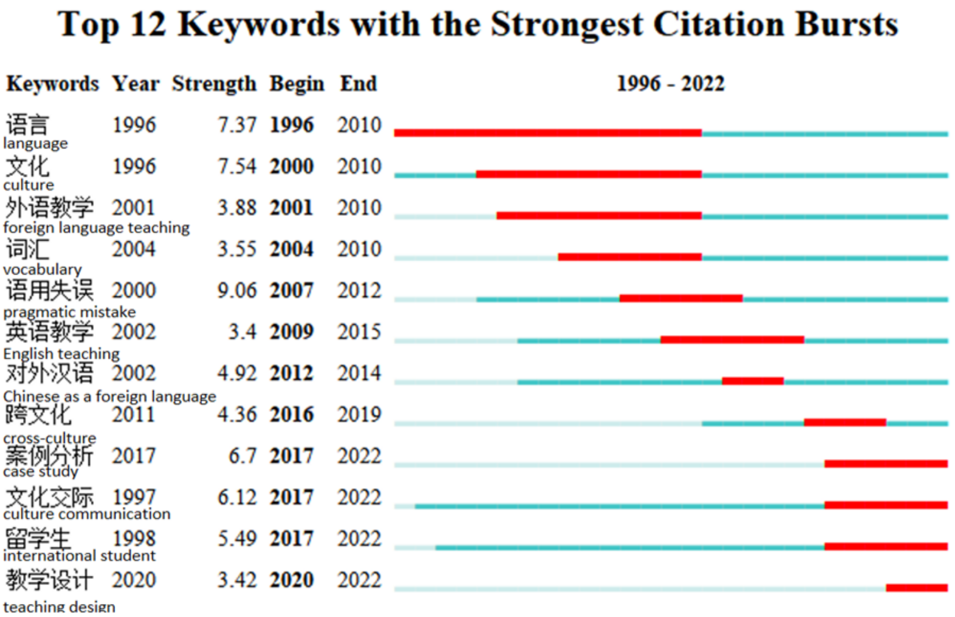


Figure 6. Analysis of Research Trends - Time Line



Top 12 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	1996 - 2022
语言 language	1996	7.37	**1996**	2010	
文化 culture	1996	7.54	**2000**	2010	
外语教学 foreign language teaching	2001	3.88	**2001**	2010	
词汇 vocabulary	2004	3.55	**2004**	2010	
语用失误 pragmatic mistake	2000	9.06	**2007**	2012	
英语教学 English teaching	2002	3.4	**2009**	2015	
对外汉语 Chinese as a foreign language	2002	4.92	**2012**	2014	
跨文化 cross-culture	2011	4.36	**2016**	2019	
案例分析 case study	2017	6.7	**2017**	2022	
文化交际 culture communication	1997	6.12	**2017**	2022	
留学生 international student	1998	5.49	**2017**	2022	
教学设计 teaching design	2020	3.42	**2020**	2022	

Figure 7.Research Trend Analysis - Keyword Burst

As can be seen from Figure 6 and Figure 7 of the clustering time line development, the overall research is carried out around the three themes of "Chinese language (or as a foreign language)", "culture (difference)", "cultural communication" and "overseas students", but the overall orientation of the research varies in different time periods. In addition, as shown in Figure 7, Year represents the year in which the keyword appears, Begin represents the year in which the keyword begins mutation, and End represents the year in which the keyword mutation is completed. Strength represents the intensity of mutation, and a large number of strength indicates a high frequency of keyword emergence (Xu & Zhao, 2021). From 1996 to 2010, the main themes were "language" "culture" "foreign language teaching" "vocabulary" "pragmatic mistakes", "Chinese as a foreign language" etc., mainly from a macro level to explore the cross-cultural communication barriers caused by cultural differences in international Chinese language education. "Language" (1996-2010, 7.37) and "culture" (2000-2010, 7.54) have the tendency of early start time, high intensity and long duration. "Pragmatic failure" (2007-2012, 9.06), among the 12 major keywords, although the mutation duration was not long, the intensity was the highest, that is, the keyword "pragmatic mistakes" appeared the most frequently during the 6-year period. From 2011 to 2022, in the past ten years, the research gradually transients from macroscopic to mesoscopic, using specific cases to analyze problems in cross-cultural communication, analyze cultural differences, and explore how to design the teaching of Chinese as a second language, a foreign language or an additional language in a cross-cultural environment. The results of data analysis indicate that "case study" and "teaching design" in cross-cultural studies are the future development trend of international Chinese language education. "Case study" (2017-2022, 6.7), "cultural communication" (2017-2022, 6.12), and "international students" (2017-2022, 5.49), the beginning time of mutation is the same as the completion time of mutation, which belong to the keywords appearing in the same period. These three keywords and "teaching design" (2020-2022, 3.42) are new research hotspots emerging in recent years, and the mutation is still continuing, indicating that these are the research hotspots of cross-cultural communication in the field of international Chinese language education.

Through the above analysis, it is found that there are many similarities with previous research results. Gao (2022) proposed that "international Chinese language education should focus on the construction of the interdisciplinary system of language and culture education, learn to remain unchanged and change, adhere to language and culture, expand cross-cultural and cross-professional development, and devote to teaching materials and national and regional studies with culture as the main line. In the training of teachers, we should face up to cultural differences, cultivate cultural consciousness, enhance cultural confidence, and expand the international perspective." His views once again emphasize the inseparable relationship between international Chinese language education and cross-cultural communication.

4. Conclusion

In this paper, CiteSpace knowledge graph software is used to make a statistical analysis of cross-cultural studies in the field of international Chinese language education based on the academic papers database of CNKI in the past 30 years (1994-2022) in China. Through sorting out and summarizing the research results, it is found that the period from 1994 to 2007 is the embryonic stage, and the related research increases slowly. The period from 2008 to 2021 is a booming development period, with the number of relevant research surging. A rapid decline began in 2021, but there was a new development trend in 2022, and the research trend began to rise again. According to the data, the authors of the literature are basically in the mode of "fighting alone", and a large scale of authors research groups have been formed. This result is consistent with the findings of Ouyang (2022). The data showed that the key words in the literature included 17 clusters, including "culture", "Chinese", "cultural communication", "differences" "international students", "cultural differences", "Chinese as a foreign language", "cultural connotation", "foreign language teaching" and "cross-culture", and the main cluster was the first 12. The results of the analysis of cited information in literature show that "cultural difference", "pragmatic mistakes", "Chinese as a foreign language", "cross-culture", "cultural teaching" and "cultural differences" are the most concerned topics. "Cultural communication", "international students", "case study" and "teaching design" are the research trends and hotspots in this field in the past two years.

Based on this, this paper has the following prospects for cross-cultural studies in the field of international Chinese language education:

(1) Cross-cultural communication is the cornerstone of international Chinese language education. The most prominent feature of international Chinese language education is that it is conducted under the cross-cultural background. Cross-cultural activities in international Chinese language education need to be perceived and understood from the perspective of cultural theory. The concepts of "original vision understanding", "cultural context"and "stereotype" are used to analyze the cross-cultural interaction and generation. It is proposed that the relationship between ours and others should be properly handled in international Chinese language education, and stories of both China and other countries should be told well (Wang, 2021).

(2) At present, the international research on identity and cultural identity is increasingly prominent. In the context of the international promotion of Chinese language, the Chinese cultural identity of international students has a great impact on their Chinese learning motivation. Most of the knowledge, culture and values that foreign students learn in China will continue to influence their political attitudes and living habits after they return their countries. At present, lack of cognition, single way of understanding and cultural shock are serious problems that foreign students are facing in the process of cultural infiltration. In view of this, the campus should first start to build cultural environment (Hu et al., 2020; Gao, 2022).

(3) In recent years, there are relatively few cross-cultural studies in online international Chinese teaching. The outbreak of COVID-19 has had a great impact on international Chinese language education and brought great challenges to international Chinese language teaching (Cui, 2022a; Cui, 2022b), forcing international Chinese language teaching to be carried out online. Although the teaching mode has been transferred to online, the cross-cultural problems of Chinese language in international Chinese language education have still not disappeared, but there are relatively few research related to cross-cultural studies in the field of international Chinese language education. Therefore, future studies should pay attention to cross-cultural communication in online teaching mode, which is conducive to the training and guidance of international Chinese language teaching and helps teachers formulate teaching and classroom management strategies, so as to better cope with cross-cultural communication in online Chinese language teaching.

Chunyan Ma is a PhD candidate in Education (Specializing in TCSOL) at the New Era University College, Malaysia. Her research interests include Teaching Chinese to Speakers of Other Languages (TCSOL), second language acquisition (SLA), cross-culture communication, learning emotion, and positive psychology.
Email: 1216551708@qq.com

Shuang Zhang is a PhD candidate in Chinese Studies at the New Era University College, Malaysia. She is currently associate professor of Chinese language education at the School of Humanities, Wuhan University of Engineering Science, China. Her research interests include Teaching Chinese to Speakers of Other Languages (TCSOL), and applied linguistics.
Email: 470046501@qq.com

Ruixin Wang (corresponding author) is currently associate professor at the Department of Education, the New Era University College, Malaysia. Her research interests include Teaching Chinese to Speakers of Other Languages (TCSOL), and Overseas Chinese Education.
Email: wangruixin527@163.com

Ting Yi is a PhD candidate in Linguistics and Applied Linguistics at Beijing Language and Culture University, China. Her research interests include Teaching Chinese to Speakers of Other Languages (TCSOL), Chinese language teaching technologies, and applied linguistics.
Email: eating327334@163.com

Chili Li (corresponding author) holds a doctoral degree in applied linguistics. He is currently associate professor of applied linguistics at the School of Foreign Languages, Hubei University of Technology, China. His research interests include Teaching Chinese to Speakers of Other Languages (TCSOL), L2 motivation, learning strategies, learner beliefs, socio-cultural theory, and others related to second language acquisition and language teaching. His works appear in System, Frontiers in Psychology, Chinese Journal of Applied Linguistics, and Springer Nature.
Email: lichili@hbut.edu.cn

Acknowledgement

This paper is supported by The Project of Philosophy and Social Science Research in Colleges and Universities of Hubei Provincial Department of Education (21G165).

References

Chen, C. M. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science & Technology*, 57(3), 359-377.

Chen, Y., Chen, C.M., Liu, Z. Y., Hu, Z. G., & Wang, X.W. (2015). CiteSpacezhishitupu de fangfalungongneng [The methodological function of CiteSpace mapping knowledge domains]. *Kexuexueyanjiu*, 2, 242-253. https://doi.10.16192/j.cnki.1003-2053.2015.02.009

Chu, Y. G., & Zhang, Z. Z. (2019). JiyuHistcite he VOSviewer de LDA moxingkeshihuafenxi [Visual analysis of LDA model based on Histcite and VOSviewer]. *Hulunbuirxueyuanxuebao*, 4, 100-108.

Confucius Institute Headquarters (Hanban). (2019). *International Curriculum for Chinese Language Education* [Guojihanyujiaoxuetongyongkechengdagang]. Beijing yuyandaxuechubanshe.

Cui, X. L. (2022a). Xinshidaiguojizhongwenjiaoyumianlinxindeketi [International Chinese education in the new era is facing new problems]. *Yunnan shifandaxuexuebao (zhexueshehuikexueban)*, 5, 46-47.

Cui, X. L. (2022b). Shijiegejujuliebianhuabeijingxia de guojizhongwenjiaoyu [International Chinese education in the context of drastic changes in the world pattern]. *Tianjin shifandaxuexuebao (shehuikexue ban)*, 4, 23-29.

Fu, J., & Ding, J. D. (2019). CiteSpace he VOSviewerruanjian de keshihuayuanlibijiao [Comparison of visualization principles between CiteSpace and VOSviewer software]. *Yanjiulunwen*, 10, 31-37.https://doi.10.13998/j.cnki.issn1002-1248.2019.10.19-0776

Gao, L. P. (2022). Lunguojizhongwenjiaoyuyuyan he wenhua de shuangchongjiangou [On the dual construction of language and culture in international Chinese education]. *Shanghai jiaotongdaxuexuebao (zhexueshehuikexue ban)*, 6, 41-50.https://doi.10.13806/j.cnki.issn10087095.2022.03.004

Geng, Z., & Gao, Y. (2022). Jinshinianguoneishuzihuahanyujiaoxueyanjiuzongshu --jiyuCiteSpace de keshihuafenxi [Summary of domestic research on digital Chinese teaching in the past decade -- Visual analysis based on CiteSpace]. *Shuzihuaguojizhongwenjiayu*, 36-42.

Gong, Y., Lai, C., & Gao, X. (2020a). The teaching and learning of Chinese as a second or foreign language: the current situation and future directions. Frontiers in Education in China 15, 1–13.

Hou, J. H., & Hu, Z. G. (2013). CiteSpaceruanjianyingongyanjiu de huiguyuzhanwang [Review and prospect of CiteSpace software application research]. *Xiandaiqingbao*, 4, 99-103.

Hu, C. H., Chen, W., & Zhao, Q. Q. (2021).

Hanyuguojituiguangshiyuxiazengqiangliuxueshengzhonghuawenhuarentong de lujingyanjiu [A Study on the Ways to Enhance the Chinese Cultural Identity of International Students from the Perspective of Chinese International Promotion]. *Wenhuachuangxinbijiaoyanjiu*, 5, 194-198.

Li, C., Chen, L., Ma, C. Y., Zhang, S., & Huang, H. Q. (2021). Strategy use among Chinese as second language learners in Mainland China from the mediationtheory perspective. *Frontiers in Psychology*,12, 1-12.https://doi.10.3389/fpsyg.2021.752084

Li, C., Yi, T., Zhang, S., Ma, C. Y., & Liu, H. G. (2022). Developing and validating a scale for measuring pre-service Chinese as an additional language teacher beliefs. *Frontiers in Psychology*12, 1-12.https://doi.10.3389/fpsyg.2022.989581

Li, W. D., Gong, Y., & Wang, Y. W. (2022). Guojizhongwenjiaoxue de shehuiwenhuashijiao: SSCI qikanwenxiankeshihuafenxi [The Socio-cultural Perspective of International Chinese Teaching: A Visual Analysis of SSCI Periodical Literature]. *Guojihanyujiaoxueyanjiu*, 2, 67-77.

Li, Y. M., & Zhai, Y. (2021). Laihualiuxueshenghanyujiaoyu 70 nian: huiguyuzhanwang [70 years of Chinese language education in China: review and prospect]. *Yuyanjiaoxueyuyanjiu*, 4, 1-10.

Liao, S. J. (2011). KexuezhishitupuhuizhigongjuVOSvieweryuCiteSpace de bijiaoyanjiu [A Comparative Study between VOSviewer and CiteSpace]. *Keji qingbaokaifayujingji*, 7, 137-139.

Ma, C., Zhang, S., & Li, C. (2022). Commentary: Chinese as a Second Language Multilinguals' Speech Competence and Speech Performance. *Frontiers in Psychology*, 13, 1-3.https://doi.10.3389/fpsyg.2022.935909

Ou Yang, L. (2022). A study of intercultural communication from the perspective of bibliometrics [Wenxianjiliangxueshijiaoxiakuawenhuajiaojiyanjiu]. *Science & Technology Information*, 135-139. https://doi.10.16661/j.cnki.1672-3791.2201-5042-3635

Peng, R. Z., Zhu, C. G., & Wu, W. P. (2021). Visualizing the Emerging Trends of Acculturation Research: A Bibliometric Analysis from 2000 to 2020. *Journal of Cross-Cultural Psychology*, 794-800.https://doi.10.1177/00220221211044782

Song, X. F., & Chi, P. J. (2016). VOSvieweryuCiteSpaceyingyongbijiaoyanjiu [Comparative Study of VOSviewer and CiteSpace Applications]. *Qingbaokexue*, 7, 108-112, 146.https://doi.10.13833/j.cnki.is.2016.07.021

Van Eck, N. J.,& Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, 111(2), 1053-1070.https://doi.10.1007/s11192-017-2300-7

Wang, H. (2021a). Xinguanfeiyanyiqingyingxiangxiayuyanyingyongyanjiu de wudarenwu [Five tasks of language application research under the influence of the COVID-19]. *Qingnianjizhe*, 9, 5.

Wang, H. (2021b). Xinguanyiqingyingxiangxia de guojizhongwenjiaoyu: wentiyuduice[International Chinese language education under the influence of COVID-19 epidemic: problems and countermeasures].*Yuyanjiaoxueyuyanjiu*, 4, 11-22.

Wang, Y. L. (2021). Guojizhongwenjiaoyuzhong de kuawenhuahudongyushengcheng [Cross-cultural interaction and generation in international Chinese education]. *Yuwenjiaoxuetongxun*, 9, 83-85.https://doi.10.13525/j.cnki.bclt.202109030

Xu, Y. R., & Zhao, L. P. (2021). Jinshinianlaiguoneizhiyejiaoyuzhengceyanjiu de redianyuzhanwang [Hot spots and prospects of domestic vocational education policy research in the past decade]. *China Adult Education*, 11, 3-11.

Zhang, L., Zhao, X., & Ye, Y. (2011). XinxikeshihuaruanjianCiteSpaceyuVOSviewer de yingyongbijiao [Application comparison of information visualization software CiteSpace and VOSviewer]. *Xinxiziyuanguanlixuebao*, 1, 95-98.

Journal of Language Teaching, *3*(3), 19-25, 2023
https://doi.org/10.54475/jlt.2023.006

Review

A critical understanding of second language acquisition from two sociolinguistic strands: The Variationist Approach and the Investment Perspective

Chang Liu

The University of Cambridge, Cambridgeshire, United Kingdom

Received: January 31, 2023 / Accepted: February 20, 2023 / Published Online: March 1, 2023
© Pioneer Publications LTD 2023

Abstract

This review explores how sociolinguistics expands our understanding of second language acquisition (SLA) by drawing upon two typical sociolinguistic strands: the variationist approach and the investment perspective. Accordingly, two empirical studies are used to illustrate the contributions of each strand, with Han's (2019) study adopting a variationist approach and Sung's (2020) study taking an investment perspective. Through a critical analysis, this paper argues that both theoretical strands contribute to the "social turn" of SLA by providing different insights into the social shaping of L2 knowledge and learning, as well as the interplay between identity construction and L2 learning. Implications, limitations, and suggestions for future studies are discussed at the end.

Keywords second language learning; sociolinguistic theory; variationist approach; investment perspective; identity

1. Introduction

Learning a second language (L2), or any language in addition to the learners' native language(s) (Block, 2003), is traditionally believed to be a context-neutral undertaking situated within learners' minds (Zuengler & Miller, 2006). Over the past three decades, however, the predominance of this cognitive paradigm in second language acquisition (SLA)[1] has been extensively challenged by socially positioned critiques (e.g., Block, 2003; Hall, 1995; Pavlenko, 2002). A prominent example driving this ongoing "social turn" (Block, 2003, p. 1) is Firth and Wagner's (1997) seminal paper, which critiques the hegemony of the cognitive-driven approach to SLA. Instead, they called for the field to become "more theoretically and methodologically balanced" (p. 286) and for a reconceptualisation of L2 learning as "emically" and "interactionally attuned" (p. 296) — a perspective embraced by many current socially-oriented theories.

Positioned within this social-cognitive debate, this paper aims to explore the theoretical insights into L2 learning contributed by sociolinguistics, a socially-situated branch concerned with exploring the "relations between the use of language and the social structure in which the language users live" (Zhang & Wang, 2016, p. 830). As a diverse and changing field of applied linguistics, sociolinguistics has adopted multiple theoretical strands to theorise L2 learning. These include but are not limited to: the variationist approach (Labov, 1963), the language socialisation theory (Duff, 1995), the theory of communities of practice (Lave & Wenger, 1991), and the investment perspective (Norton Peirce, 1995). While the variationist approach traditionally investigates variability in learner language, the remaining strands focus on the dynamic social processes of L2 learning.

The current review draws on the variationist approach and the investment perspective to discuss how sociolinguistics expands our understanding of SLA. Pioneered by Labov (1963), the variationist approach primarily employs quantitative research means to investigate the causes of socially-patterned variations in language use, which is the fundamental concern of sociolinguistics (Holmes & Wilson, 2017). In contrast, the investment perspective, proposed by Norton Peirce (1995) and expanded by Darvin and

---

[1] This paper uses the terms L2 learning and SLA interchangeably to refer to both the scholarly field of inquiry in applied linguistics and the process of learning an additional language after the successful acquisition of the L1(s), or first language(s) (Ortega, 2011).

Norton (2015), adopts qualitative paradigms to explore the social structures and power relations imbued in the L2 learning process, an angle unexamined by the variationist tradition. Therefore, given the fundamental differences between these two approaches in terms of origins, theoretical foci, and research methods, valuable insights about L2 learning through a sociolinguistic lens could be derived from a critical discussion of both.

Accordingly, two empirical studies will be used to illustrate the contributions of each strand. While Han's (2019) paper investigated L2 sociopragmatic performance from a variationist sociolinguistic perspective, Sung's (2020) study examined learners' investment in L2-mediated social interactions through an investment model. These two papers are selected as they tellingly reflect the core contributions of the chosen sociolinguistic strands, which transcend beyond the traditional cognitive awareness of SLA. Specifically, the reviews found that the variationist approach deepens our understanding of SLA by revealing its socially-conditioned nature, whereas the investment perspective contributes to this field by uncovering the value-laden and power-imbued SLA process. Moreover, both strands illustrate the interplay between L2 learning and identity constructions, with the investment perspective exploring the influence of social constraints and learner agencies on identity negotiations in greater depth.

The following sections discuss how each theoretical approach contributes to insights about SLA according to major themes (i.e., social shaping of L2 knowledge and learning; the interplay between identity construction and L2 learning) that emerged from the chosen empirical studies. Implications, limitations, and directions for future research will be offered at the end of this review.

2. Social Shaping of L2 Knowledge[2] and Learning

Stemming from the sociolinguistic tradition, the variationist approach argues for the systematicity of variability in patterns of language use (Geeslin, 2020), which means that the differences in linguistic features produced by L2 learners can be explained by a range of internal (i.e., linguistic and developmental factors) and external elements (i.e., social factors) (Romaine, 2003). The latter, which involve the social context, interlocutors, and L2 learners' social categories, constitute the primary focus of variationist sociolinguistics (Regan, 2013). Unlike the cognitive framework that perceives L2 learning as merely the development of grammatical competence (Firth & Wagner, 1997), variationist sociolinguistics concerns with the acquisition of L2 forms in socially appropriate ways, i.e., sociopragmatic ability (Regan, 2013).

Drawing on a variationist perspective, Han (2019) focused on L2 Chinese learners' sociolinguistic pragmatic performance by exploring their use of sentence-final particles (SFPs)[3] in non-interrogative sentences. Methodologically, the data for this study came from the conversations of eight L2 users who appeared on the popular Chinese talk show *Informal Talks*. Apart from the speech data, Han also collected text data from the participants' Weibo, a Chinese social media platform. Having identified a range of linguistic and social variables informed by existing literature, Han used Rbrul to quantitatively determine the ones that influenced the varied presence of SFPs, followed by qualitative case analyses of how participants utilised SFPs in different situations and how these were perceived by others.

The findings of this study illustrated the socially-conditioned character of L2 knowledge and learning. For example, based on the Rbrul analysis, Han found that the differences in time spent in China and in gender-related personality were statistically significant factors that impacted the variations in SFPs. Regarding the former, it was revealed that more time spent in China corresponded to more frequent use of SFPs. This phenomenon, according to Han, can be explained by the fact that with increased opportunities for language socialisation with native speakers, L2 learners gained a better understanding of the social indications of SFPs, and thus expressed themselves more effectively using SFPs. These findings shed light on the influences of participants' social categories (e.g., gender-related personality) and socialisation (e.g., learner-native interactions) on L2 knowledge and learning.

Another valuable insight discussed in Han's study that illustrated the socially-shaped nature of SLA is the role of L1 transfer in the acquisition of sociopragmatic competence. Unlike the traditional contrastive analysis that perceives negative transfers of L1 as causes of potential L2 errors (Al-khresheh, 2016), this study showed that the role of L1 was a non-significant factor in influencing the acquisition of SFPs. While this result could have been affected by the relatively small corpus size, it was Han's explanation that cast new light on the nature of SLA. She attributed the minimal effect of L1 to the fact that the subjects' motivation to be "socialised into active agents outweighs linguistic difficulty of SFPs" (p. 58). This explanation implied that the positive social influence on L2 learning could potentially compensate for negative L1 transfers, thus confirming the role played by social factors in SLA.

Besides the impact of social variables on linguistic variations, L2 users' social intentions and possible reactions of the audience were also found to shape L2 knowledge and learning. In the case of Daddy Mu[4], certain SFPs were used to enact an amiable personality to his audience, particularly on Weibo. For example, with the social intention of reminding his fans to watch his show, he added the SFP "o (哦)" at the end of the imperative sentence ("you must watch!" [p. 61]) to create a friendly and casual mood. Although Daddy Mu mostly projected a masculine image, he had learnt to manipulate SFPs to convey expressive feminine traits during interactions with his fan groups. As argued by Han, this meaningful speaker-listener relationship facilitated the learning aspects of SFPs, which in turn broadened the audience base of the L2 user. In this sense, the variationist approach demonstrates how L2 knowledge and learning can be shaped by learners' social intentions and the possible perceptions of the audience, again exemplifying the social dimension of SLA.

---

[2] L2 knowledge is also known as interlanguage, which is defined as the type of language generated by the L2 learner that shares features of both the learner's L1 and the target language (Selinker, 1972).

[3] In Mandarin Chinese, most SFPs are grammatically optional morphemes typically attached to the end of a statement or question (Wamsley, 2019) to imply the speaker's attitude, level of assertiveness, and evidentiality (Simpson, 2014). The L2 acquisition of Chinese SFPs is said to be particularly difficult due to the complexity of the rules and the insufficient instruction of such rules in formal teaching (Han, 2019).

[4] "Daddy Mu" is a nickname given by the audience to Mohammed, a

While variationist sociolinguistics is concerned with socially-mediated language variations, the investment perspective extends this line by delving into the complex socialisation processes of L2 learning. The sociological construct of investment was proposed by Norton Peirce (1995) as a complement to the psychological concept of motivation in order to explain the non-linear L2 learning phenomenon, where deeply motivated learners may still refuse to participate in learning opportunities under socially hostile circumstances. Two decades later, Darvin and Norton (2015) developed the model of investment as a response to the new world system "characterized by mobility, fluidity, and diversity" (p. 51). This expanded model highlights three intersecting factors (i.e., ideology, capital, and identity) that dynamically constitute the complexities involved in L2 learning. Core to this model lies the belief that learners invest in an L2 with the understanding that they will be rewarded with myriad material and symbolic resources, which will in turn translate to their enhanced social status (Darvin & Norton, 2017).

Similar to variationist sociolinguistics, the investment perspective focuses on learners' sociolinguistic ability rather than their grammatical competence. This is suggested by Darvin and Norton (2015), who construe individuals' forms of linguistic capital as "styles and registers", which are "measured against a value system that reflects the biases and assumptions of the larger sociocultural context" (p. 45). What makes the two theoretical strands differ is that the investment perspective tends to treat the linguistic and even nonlinguistic resources as part of the L2 learners' capital. This coincides with Darvin and Norton's (2018) argument that learners are able to "assemble and engage more complex linguistic and non-linguistic repertoires, where [L2] becomes just one of many resources" (p. 4). This means that the investment perspective views L2 knowledge as an integral part of one's entire linguistic repertoire.

Grounded in the investment perspective, Sung (2020) investigated a group of mainland Chinese university students' investment in Cantonese learning during their cross-border studies in Hong Kong. Specifically, this study focused on participants' negotiations of identities, forms of capital, as well as the impact of ideologies on their L2 investment. Based on a thematic analysis of interview data from a larger project involving 21 mainland Chinese university students, Sung identified the dynamic interactions of identity, capital, and ideology in participants' Cantonese learning experiences, which reflected the multilayered L2 learning process conditioned by hierarchical power relations and preconceived values.

One example was participants' struggles in converting existing resources, such as their L1, as affordances for L2-mediated interactions. The investment perspective conceives L1 as a form of linguistic capital, which can potentially create more L2 learning opportunities. In Sung's study, however, participants experienced difficulties transforming their proficiency in Mandarin, a form of linguistic capital they already possessed, into opportunities for Cantonese learning. This was because while Mandarin was deemed as a highly valued type of capital in mainland China, it was regarded as "a peripheral language" (p. 11) in the Hong Kong university context. As the locals rarely sought opportunities to practise Mandarin with the participants, proficiency in Mandarin could not be usefully capitalised to access L2-mediated social opportunities. This finding suggested that, when entering a new social space, learners' linguistic capital can be devalued by those in power with a predetermined value system.

The role of ideologies also illustrates how dominant views of the powerful could influence L2 learning. The incorporation of ideology in this model was enlightened by De Costa's (2010) call for an explicit naming of ideology in SLA in order to render systemic patterns of control visible. Darvin and Norton (2015) conceptualise ideology in broader terms beyond the dimensions of governance or language, as normative beliefs that frame societies and decide ways of inclusion and exclusion. In Sung's study, for example, anti-mainlandisation ideology due to the tense Hong Kong-Mainland political relationship and the deep-seated negative stereotypes towards mainlanders made it difficult for participants to gain meaningful Cantonese learning opportunities. This finding indicated that L2 learning experiences are mediated by ideologies that shape the assumptions, values, and beliefs held by the more powerful others towards L2 learners.

In summary, both studies have provided ample evidence to the contributions of sociolinguistic theories to SLA by unveiling the social shaping of L2 learning and knowledge. Han's study not only showed the systematic nature of linguistic variations and their correlations with social categories but also uncovered how these variations and the acquisition of sociopragmatic ability were conditioned by L2 learners'

---

journalist from Egypt, due to many of his macho opinions. He possessed the lowest frequency of the use of SFPs (17%).

social intentions and their audiences' potential responses. Meanwhile, Sung's study demonstrated the situated nature of L2 learners' linguistic resources and the value-laden, ideology-infused L2 learning experiences. However, the discussions so far have barely analysed learners' agency during L2 learning. This will be further explained in the next section on the interplay between identity construction and L2 learning to enrich the understanding of SLA from a sociolinguistic perspective.

3. The Interplay between Identity Construction and L2 Learning

The construct of identity has been a focus of research in variationist sociolinguistics ever since its inception in the early 1960s (Labov, 1963). While different variationist traditions disagree on the definition and the role of identity in language variation, this paper focuses on the third wave[5] of variationist sociolinguistics (TWVS). As a relatively recent approach to the examination of sociolinguistic variation (Eckert, 2016), TWVS is most relevant to Han's (2019) study of participants' identity construction through language variation. While the first two waves concern static groups of speakers and associate identity with category affiliation, TWVS focuses on social meaning and speaker agency (Eckert, 2012). It treats language variation as an expression of social identities by speakers through stylistic practices (Eckert, 2012). Identities, or senses of self (Duff, 2013), are regarded as socially constructed, dynamic, and changeable (Drummond & Schleef, 2016), challenging the essentialised view of identity as binary and stable from a conventional cognitivist strand (Davies, 1991).

Instead of examining the relationship between language variation and the biological category of sex, Han examined gender-associated variations in Chinese language use, as evidenced by the self-presentations of gender-related personality characteristics. Individuals exhibit varying levels of a combination of both masculine (e.g., independence and assertiveness) and feminine characteristics (e.g., sensitivity and compassion). Through the example of Madam Qian's[6] identity performance, Han illustrated L2 learners' strategic use of SFPs to construct their desired identities. The qualitative analysis showed that Madam Qian frequently manipulated SFPs to evoke a warm and sympathetic persona. For instance, during the debate on whether we should offer seats to the elderly on public transport, Madam Qian skillfully drew on a variety of SFPs, including ya (呀), a (啊), and ma (嘛), to emphasise his affective attitudes and construct his public self-image. This exemplifies that the acquisition of SFPs could empower L2 learners by granting them social agency as users of this language to articulate gender-related social identities.

Another example that showed the transformative role of L2 learning in constructing identities is Madam Qian's choice of wearing a female hat when presented with traditional hats from different cultures in the TV show. While his choice was perceived by fellow participants as looking like a female, Madam Qian responded by manipulating a series of feminine SFPs, such as ma (嘛), to soften his responses. His language further strengthened the incoherence between his biological sex and sociological gender identity at that moment. Notwithstanding this incongruence, his gendered use of SFPs was appreciated and well-received by other participants and his fans on social media, which, in turn, broadened his audience base. This instance again demonstrated how successful L2 acquisition grants learners more agency to construct desired identities, potentially creating more favourable L2 social opportunities.

Similar to TWVS, the investment perspective is also grounded in the poststructuralist perspective, which views identity as multiple, shifting, contradictory, and socially constructed (Norton, 2000). Specifically, identity is conceptualised as "the way a person understands his or her relationship to the world, how that relationship is constructed across time and space, and how the person understands possibilities for the future" (Norton, 2013, p. 4). In Sung's (2020) study, while participants performed their agency to construct identities as Cantonese communicators so as to access more Cantonese-mediated interactions, the challenges they experienced hindered their L2 learning. Drawing on Miller's (2004) concept of *audibility*, or "the degree to which speakers sound like, and are legitimated by users of the dominant discourse" (p. 291), Sung revealed that participants struggled to gain audibility during class discussions due to their limited Cantonese proficiency, which subsequently discouraged them from continuing speaking. Based on the findings, this lack of recognition as legitimate Cantonese communicators by the locals also made it difficult for participants to claim in-group membership, thus limiting their L2 learning experiences.

While opportunities for L2 learning are constrained by identities imposed by others, Sung's study also showed how agents' self-positioning could create obstacles in constructing desired identities, which further affected their L2 learning. For example, standard language ideology, which in this study referred to the importance of speaking Cantonese with a standard accent, was internalised by participants. Although the participants aspired to speak Cantonese with the ideal accent, they still struggled to do so and suffered "a sense of linguistic inferiority" (p. 10). This "negative self-positioning" (p. 11) undermined their desired identities as authentic learners and users of Cantonese, which further dampened their confidence to engage in L2-mediated socialisations.

---

[5] Variationist sociolinguistics has come in three waves of traditions. The first focuses on documenting relationships between linguistic variables and macrosocial categories spanning large populations. The second wave employs ethnographic methods to examine the relation between variation and local social categories (Eckert, 2016).

[6] "Madam Qian" is a nickname used by fans for James, a kindergarten teacher from Nigeria, as he is very emotional when expressing his opinions on the show. He possessed the highest frequency of use of SFPs (33%).

Another type of self-positioning relates to learners' pursuit of imagined identities, or their ideal sense of selves, affiliations, and social groups they aspire to be part of (Kanno & Norton, 2003). The crucial relationship between imagination and identity is highlighted by Wenger (1998), who conceptualises imagination as "a process of expanding our self by transcending our time and space and creating new images of the world and ourselves" (p. 176). As identified from the data, participants faced dilemmas between investing in Cantonese and English as they were uncertain about their imagined communities, known as "groups of people not immediately tangible and accessible with whom we connect through the power of imagination" (Kanno & Norton, 2003, p. 241). Some participants placed more emphasis on investing in English as they imagined themselves being able to function in a global educational or workplace context. Thus, their devotion to learning Cantonese was not just associated with their desire to establish identities as legitimate speakers, but also related to how they position themselves in the future. On this matter, ambivalence about imagined identities could also affect L2 learners' agency to invest actively in their L2 practices.

To sum up, both studies illustrate the "language-identity nexus" (Joseph, 2004, p. 12) — the mutual shaping and reinforcement between identity construction and language learning. While the variationist approach highlights the empowering role of L2 acquisition in identity constructions, the investment perspective illustrates how L2 learners' identity performance can be circumscribed by power differentials, how L2 learners contest this power imbalance, as well as how learners' self-positioning affects L2 learning.

4. Conclusions

This review has discussed how sociolinguistic theory contributes to insights about L2 learning from the perspectives of two theoretical strands — the variationist approach and the investment perspective through the illustrations of two empirical studies. It argues that both theoretical approaches provide valuable insights into SLA by revealing the social shaping of L2 knowledge and learning, as well as the interplay between identity construction and L2 learning, albeit to different degrees and from different angles. The variationist approach confirms the relevance of social factors to L2 learning and highlights the transformative role of SLA in identity constructions, which may translate to greater L2 social interactions. The investment perspective takes a more radical approach by uncovering the power-laden nature of L2 learning and identity negotiations, along with L2 learners' agency, in confronting social constraints. Therefore, these two strands could be visualised as occupying different positions on the spectrum of sociolinguistics, with the variationist approach situated at the initial part and the investment perspective at the extreme end.

The findings of this paper could be used to inform possible reasons for successful and unsuccessful SLA from the perspective of sociolinguistics, which will shed light on pedagogies and policies regarding L2 education. For example, in regards to the social constraints on L2 learners' identity performance, teachers and policymakers could encourage more supportive "audiences" and "interlocutors" (e.g., local students in the study abroad context and local people interacting with L2 learners) and empower learners by promoting the value of their L1. Moreover, lessons on hidden power relations and social structures could be provided not just to L2 learners but to all students living in culturally and linguistically diverse settings to raise their awareness of the social mechanisms underpinning L2 learning. Lastly, sociopragmatic competence should be given more emphasis in L2 education so as to increase learners' chances of exercising agency to access L2 learning opportunities.

Despite an in-depth analysis of the contributions made by two sociolinguistic approaches, this review is not free from limitations. For example, aside from the areas of contributions discussed, other aspects, such as the context of SLA and research methodologies, are beyond the scope of this review. Moreover, the arguments about the two theoretical strands are mostly based on the two empirical studies, which might be limited in offering insights into SLA. Therefore, future research could seek to address these limitations by exploring other areas of concern and drawing on a broader range of empirical studies.

Chang Liu is a graduate of the University of Cambridge, where she studied MPhil Research in Second Language Education. Prior to her master's studies, she completed a BA in Education Studies from the University College London. Currently, her research interest lies in language investment, informal language learning, digital literacies, and social inequalities.
Email: cl868@cantab.ac.uk

References

Al-khresheh, M. H. (2016). A review study of contrastive analysis theory. *Journal of Advances in Humanities and Social Sciences*, *2*(6), 330-338. https://doi.org/10.20474/jahss-2.6.5

Block, D. (2003). *The social turn in second language acquisition*. Georgetown University Press, 3240 Prospect Street, NW, Washington, DC 20007.

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

Darvin, R., & Norton, B. (2017). Language, identity, and investment in the twenty-first century. *Language policy and political issues in* *education*, 1-15. https://doi.org/10.1007/978-3-319-02320-5_18-2

Darvin, R., & Norton, B. (2018). Identity, investment, and TESOL. In J. I. Liontas (Ed.), *TESOL encyclopaedia of English language teaching*. https://doi.org/10.1002/ 9781118784235.eelt0802

De Costa, P. I. (2010). Reconceptualizing language, language learning, and the adolescent immigrant language learner in the age of postmodern globalization. *Language and Linguistics Compass, 4*(9), 769-781. https://doi.org/10.1111/j.1749-818X.2010.00229.x

Drummond, R., & Schleef, E. (2016). Identity in variationist sociolinguistics. In *The Routledge handbook of language and identity* (pp. 76-91). Routledge. https://doi.org/10.4324/9781315669816.ch3

Duff, P. A. (1995). An ethnography of communication in immersion classrooms in Hungary. *Tesol Quarterly, 29*(3), 505-537. https://doi.org/10.2307/3588073

Duff, P. A. (2013). Identity, agency, and second language acquisition. In *The Routledge handbook of second language acquisition* (pp. 428-444). Routledge.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology, 41*, 87-100. https://doi.org/10.1146/annurev-anthro-092611-145828

Eckert, P. (2016). Third wave variationism. In *The Oxford Handbook of Sociolinguistics (Oxford: Oxford University Press)*. https://doi.org/10.1093/oxfordhb/9780199935345.013.2

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The modern language journal, 81*(3), 285-300. https://doi.org/10.1111/j.1540-4781.1997.tb05480.x

Geeslin, K. (2020). Variationist perspective(s) on interlocutor individual differences. In L. Gurzynski-Weiss (Ed.), *Cross-theoretical explorations of interlocutors and their individual differences* (pp. 127-158). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.53.06gee

Hall, J. K. (1995). (Re)creating our worlds with words: A sociohistorical perspective of face-to-face interaction. *Applied linguistics, 16*(2), 206-232. https://doi.org/10.1093/applin/16.2.206

Han, Y. (2019). Chinese L2 users as active social agents: Sentence final particles variation and identities. *Applied Pragmatics, 1*(1), 46-67. https://doi.org/10.1075/ap.00003.han

Holmes, J., & Wilson, N. (2017). *An Introduction to Sociolinguistics* (5th ed.). Routledge. https://doi.org/10.4324/9781315728438

Kanno, Y., & Norton, B. (2003). Imagined communities and educational possibilities: Introduction. *Journal of language, identity, and education, 2*(4), 241-249. https://doi.org/10.1207/S15327701JLIE0204_1

Joseph, J. (2004). *Language and identity: National, ethnic, religious*. Basingstoke: Palgrave Macmillan. https://doi.org/10.2104/aral0728

Labov, W. (1963). The social motivation of a sound change. *Word: Journal of the International Linguistic Association, 19*(3). 273-309. https://doi.org/10.1080/00437956.1963.11659799

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press. http://hdl.handle.net/10125/44675

Miller, J. (2004). 10. Identity and language use: The politics of speaking ESL in schools. In *Negotiation of identities in multilingual contexts* (pp. 290-315). Multilingual Matters. https://doi.org/10.21832/9781853596483-013

Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. Longman Publishing Group. https://www.academia.edu/14417937/Identity_and_language_learning_Gender_ethnicity_and_educational_change

Norton, B. (2013). *Identity and language learning*. Multilingual matters. https://doi.org/10.21832/9781783090563

Norton Peirce, B. (1995). Social identity, investment, and language learning. *TESOL quarterly, 29*(1), 9-31. https://doi.org/10.2307/3587803

Ortega, L. (2011). *Second language acquisition* (pp. 191-204). Routledge.

Pavlenko, A. (2002). Poststructuralist approaches to the study of social factors in second language learning and use. In V. Cook (Ed.), *Portraits of the L2 User* (pp. 275-302). Multilingual Matters.

Regan, V. (2013). Variation. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge Handbook for Second Language Acquisition*. Cambridge University Press. http://hdl.handle.net/10197/4698

Romaine, S. (2003). Variation. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 409-435). Oxford: Blackwell Publishing.

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*(1-4), 209-232. https://doi.org/10.1515/iral.1972.10.1-4.209

Simpson, A. (2014). Sentence-Final Particles. In C.T. J. Huang, Y.H.A. Li, & A. Simpson (Eds.), *The handbook of Chinese linguistics* (pp. 156-179). New York: Wiley. https://doi.org/10.1002/9781118584552.ch7

Sung, C. C. M. (2020). Cantonese learning, investments, and identities: Mainland Chinese university students' experiences during cross-border studies in Hong Kong. *Learning, Culture and Social Interaction, 26*, 100415.

https://doi.org/10.1016/j.lcsi.2020.100415

Wamsley, J. C. (2019). Gendered Usage of Sentence-Final Particles in Mandarin Chinese. *IULC Working Papers, 19*(1).

Wenger, E. (1998). *Communities of practice*. Cambridge: Cambridge University Press.

Zhang, H., & Wang, N. (2016). Sociolinguistics and English teaching in China. *Theory and Practice in Language Studies*, *6*(4), 830. https://doi.org/10.17507/tpls.0604.21

Zuengler, J., & Miller, E. R. (2006). Cognitive and sociocultural perspectives: Two parallel SLA worlds?. *TESOL quarterly*, *40*(1), 35-58. https://doi.org/10.2307/402645

Journal of Language Teaching, *3*(3), 26-38, 2023
https://doi.org/10.54475/jlt.2023.007

ORIGINAL RESEARCH

A corpus-based study to evaluate the generativist explanation of children's error patterns in questions

Yiran Du

Department of English, University of Liverpool, Liverpool, United Kingdom

Received: January 29, 2023 / Accepted: February 20, 2023 / Published Online: March 1, 2023
© Pioneer Publications LTD 2023

Abstract

This study explores whether the generativist account, specifically the integration theory, could explain children's percentage of errors in questions in general and whether it also applies to yes-no and non-subject wh-question. The current study adopts a corpus-based method to compare 2-to-3-year-old children's percentages of errors in questions (and in yes-no and wh-question separately) including auxiliary DO and auxiliary HAVE. The results show that children's rate of errors in questions including auxiliary DO is higher than that including auxiliary HAVE, which is also applicable to yes-no and non-subject wh-questions. The findings indicate that the generativist theory of child language acquisition could successfully explain children's patterns of errors in questions. This study also emphasises the impact of the question type which should be carefully considered when constructing and improving the generativist theory of child question formation. The study provides empirical evidence for improving and refining the generativist account of child language acquisition generally and language question acquisition specifically.

Keywords Generativism; child language acquisition; Universal Grammar; yes-no question; wh-question; error

1. Introduction

Generativism, developed by Noam Chomsky, has been an influential approach to studying child language acquisition since the 1950s after it supplanted the behaviourist approach to exploring language behaviour (Traxler, 2016). Its assumption that children are born with innate linguistic knowledge termed Universal Grammar (UG) and the subsequent Principles and Parameters framework is widely used in language acquisition research (Kania, 2016). Particularly, children question formation which involves inversion (or movement) has attracted many researchers' interests (Santelmann et al., 2002), and their research in inversion and child question formation make great contributions for constructing and improving generative account of child question acquisition (e.g., Borer & Wexler, 1987; De Villiers, 1991; Erreich, 1984; Ingram & Tyack, 1979; Klee, 1985; Klima & Bellugi, 1966; Kuczaj, 1976; Labov & Labov, 1978; Radford, 1990, 1994; Rowland, 2007; Theakston et al., 2001, 2005; Valian, 1991). Specifically, many generativists propose that inversion or movement is an essential component of UG which is constantly available to children (e.g., De Villiers, 1991; Stromwold, 1990), and children could utilise this innate linguistic knowledge to form adult-like questions from the very beginning of the language acquisition process (Rowland, 2007; Theakston et al., 2005). However, children also produce a considerable number of questions with various errors at the same time (see Bellugi, 1965, 1971), and the error tends to show systematic patterns (Kania, 2016), which should be explained by any theory aiming to describe the process of child language acquisition (Rowland, 2007). Although different solutions are proposed by many researchers such as the maturation theory (e.g., Babyonyshev et al., 2001; Borer & Wexler, 1987, 1992; Klima & Bellugi, 1966; Radford, 1990, 1994; Vainikka, 1993) and the production limitation theory (e.g., Bloom, 1990; Valian, 1991), a more promising idea is that children gain all components of UG at birth, but they also have to acquire specific rules of inflexion system from input and integrate them with the innate knowledge of inversion or movement to form questions (e.g., Santelmann et al., 2002; Stromwold, 1990), which could explain children's systematic question error patterns. For example, Santelmann et al. (2002) predict that children will make more mistakes when producing English questions including auxiliary DO[1] than those including auxiliary HAVE[2] and modal auxiliaries because the former requires additional knowledge of English inflexional rules. However, Rowland (2007) points out that this theory only applies to yes-no questions rather than non-subject wh-question[3] by comparing children's percentage of errors in yes-no and non-subject wh-question including auxiliary DO and modal auxiliaries. Due to the controversy of whether the generativist idea could explain the error patterns in children's questions, particularly non-subject wh-question, more empirical evidence is needed and this study aims to replicate Rowland's (2007) study to compare children's percentage of errors of questions including auxiliary DO and auxiliary HAVE. This study could help construct and improve the generativist account of child language acquisition, particularly child question acquisition.

2. Literature Review

2.1. The generativist theory of child language acquisition

The generativist theory of child language acquisition (sometimes also referred to as nativism or gernerativism) started from the cognitive revolution (see Miller, 2003 for an overview) initiated by Noam Chomsky, one of the founders of cognitive science, with his work on linguistic theory and theory of language acquisition (Chomsky, 1957, 1959, 1964, 1965) together with others' influential publications (e.g., Miller, 1956; Newell et al., 1958) after the mid 20th century. Particularly, Chomsky's review (Chomsky, 1959) on representative behaviourist B. F. Skinner's *Verbal Behavior* (Skinner, 1957) challenged the foundations of behaviourism, and rejected its explanations of child language acquisition.

Behaviourism was the leading approach to studying psychology (including language behaviour) from the early 20th century to the late 1950s after it supplanted introspection as the primary paradigm to understand the cognitive abilities of humans in psychology (Traxler, 2016). The critical principle of behaviourism is that the invisible mental representations and processes in the 'black box' cannot be observed directly. Therefore, any theory in psychology can only be constructed by studying the relationships between observable external stimuli and human behaviour, and any theory appealing to invisible mental events should be abandoned (Kania, 2016; Traxler, 2016). The behaviourist tried to explain behaviour by studying how animals learn associations between stimuli (e.g., classical conditioning; Pavlov, 1927; Watson & Rayner, 1920) or situations (e.g., instrumental/operant conditioning through reward and punishment; Skinner, 1957). Specifically, Skinner (1957) claims that instrumental conditioning could also be used to explain human verbal behaviour, namely language, and children imitate caregivers' speech to acquire language because there is a reward for their speech behaviour. However, behaviourism and Skinner's explanation of child language acquisition was almost abandoned and supplanted by gernerativism due to the inadequacy of behaviourism to explain the certain phenomenon of human language behaviour (Traxler, 2016). For example, Chomsky (1959) points out that the native speaker (or even a 5-year-old child) can easily know *Colourless green ideas sleep furiously* is grammatical regardless of its strange semantics and *Sheep green colourless furiously ideas* is ungrammatical even though they never hear or encounter these sentences before, which cannot be simply explained by imitation.

On the contrary, the generativist constructs the theory of language and theory of child language acquisition by appealing to hypothetical or invisible mental representations and processes even though they are not directly observable (Kania, 2016; Traxler, 2016). The main assumption of generativism comes from the famous poverty of stimulus argument (see Chomsky, 1980, 1986), which is summarised by (Rowland, 2013) as follows: children can only acquire language through information from environment or their genetic inheritance; given that data provided by the environment is not sufficient for children to learn a language, some innate linguistic knowledge must be encoded in genes. Based on this assumption, i.e., innate linguistics knowledge in human genes, Chomsky (1965) argues that there should be an inborn language faculty, also named Language Acquisition Device, hard-wired in humans' minds. Language Acquisition Device is also regarded as a language-specific cognitive module for language development, independent of other cognitive modules (Fodor, 1983) and needs biological explanation (Garnham, 2013). This Language Acquisition Device or language-specific cognitive module contains innate linguistic knowledge, termed UG by Chomsky (1965). Because of UG, children can acquire language within a short period under the condition that there is a lack of rich linguistic input around them (Kania, 2016). Based on the hypothesis of UG, the strong continuity hypothesis claims that the theory of grammar for adults are supposed to be applied to the theory of children's grammar, which should explain their production of grammatical and ungrammatical sentences (Hyams, 1986; Pinker, 1984) while the weak continuity hypothesis argues that children do not have to mater all grammar of adults and they only need to utilise general UG principles (see next paragraph) to produce sentences (Clahsen, 1990; Haan, 1987; Jordens, 1990). However, there is almost no agreement regarding the exact content of UG (Kania, 2016). Chomsky changed his description of UG several times (Rowland, 2013) from Standard Theory (see Chomsky, 1957, 1965) to Principles and Parameters (based on Government and Binding Theory) (see Chomsky, 1981) to the Minimalist program (see Chomsky, 1995) to recursion as the narrow language faculty which is unique to humans (see Hauser et al., 2002).

---

[1] Capital letters DO refer to all its auxiliary subtypes (*do, does, did*) in this article.

[2] Capital letters HAVE refer to all its auxiliary subtypes (*have, has, had*) in this article.

[3] This study focuses on non-subject wh-questions which includes all object and adjunct non-subject wh-questions because inversion is not required in subject non-subject wh-question formation (see also the corpora section below).

Because the Principles and Parameters framework is one of the most influential generative explanations of child language acquisition and is still widely used by many language acquisition studies (including this study) (Kania, 2016), it is therefore reviewed here briefly. Children are assumed to be encoded biologically with a set of principles and parameters of linguistic knowledge which help them to acquire language (Kania, 2016; Lust, 2006; Rowland, 2013). The principles manifest themselves in all languages universally, while the parameters can set different values based on different languages (Rowland, 2013). Children set parameters of their language at the very beginning of their lives with limited exposure to their mother language (Kania, 2016). For example, Santelmann et al. (2002) claim that the knowledge of movement or inversion is a principle in UG universally available to all natural languages, and different languages can use this principle specifically, which implies a set of possible parameters. For instance, I to C movement is considered a UG parameter (Fodor & Sakas, 2004). Principles and Parameters can help explain the differences in syntactic rules in various languages and children's unexpectedly sophisticated linguistic knowledge (Rowland, 2013).

However, even though the generativist theory gains a lot of credit in explaining children's performances in language acquisition (including question acquisition), it still faces some criticisms such as the problem of poverty of stimulus argument (see empirical assessment in Pullum & Scholz, 2002), the linking problem (see Tomasello, 2005), and inadequacy to explain various errors in children's early speech (see Rowland, 2013). Therefore, more empirical evidence is required to support and improve the generativist theory of child language acquisition.

2.2. The generativist theory to explain child question acquisition

The generativist theory of child question acquisition is closely related to how the questions are formed based on the theory of generative transformational grammar (Chomsky, 1981). Four English examples concerning the topic of this study are listed below:

A) *Does LeBron James eat Taco*[4] *?* (yes-no question including auxiliary DO)

B) *Has LeBron James eaten Taco?* (yes-no question including auxiliary HAVE)

C) *What does LeBron James eat?* (non-subject wh-question including auxiliary DO)

D) *What has LeBron James eaten?* (non-subject wh-question including auxiliary HAVE)

Figure 1 presents information about the formation of A and B. A is formed from its declarative counterpart (*LeBron James eats Taco*). According to the inflexional rules (number, tense, and person), the auxiliary DO (*does*, in this case) is generated automatically in the inflexion phrase's (IP) head position and the inflexional suffix of the main verb *eats* (-s) disappears. It then moves to the head position of the complementizer phrase (CP) (so-called I to C movement) (*Does LeBron James eat Taco*) and leaves a deletion trace (x) in the initial position. Similarly, originating from B's declarative counterpart (*LeBron James has eaten Taco*), its auxiliary HAVE (*has*, in this case) moves from the head position of IP to that of CP (*Has LeBron James eaten Taco*) and leaves a deletion trace.

Figure 1. Examples of formal representation of the formation of the yes-no question including auxiliary DO and auxiliary HAVE

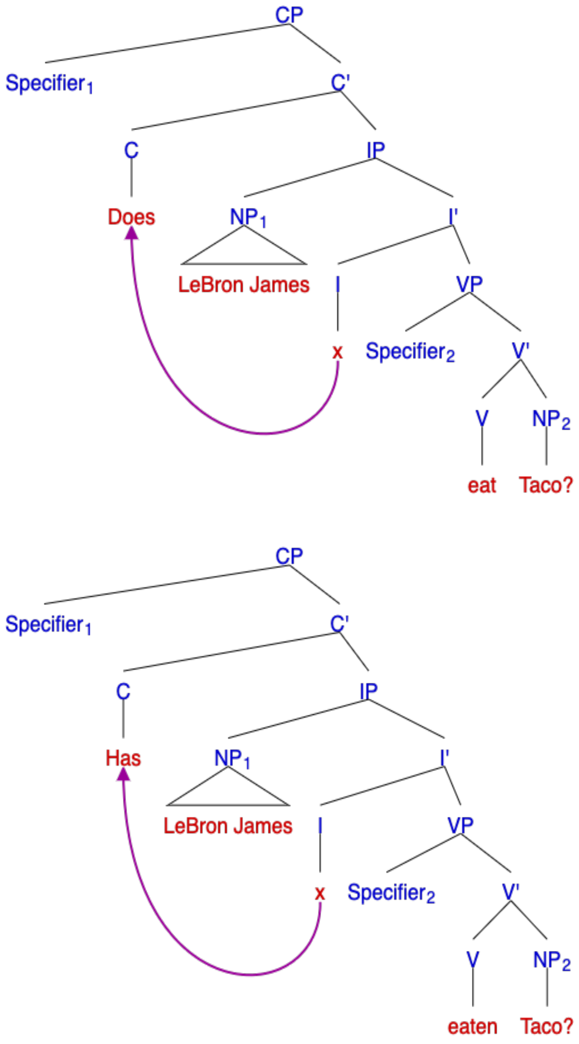
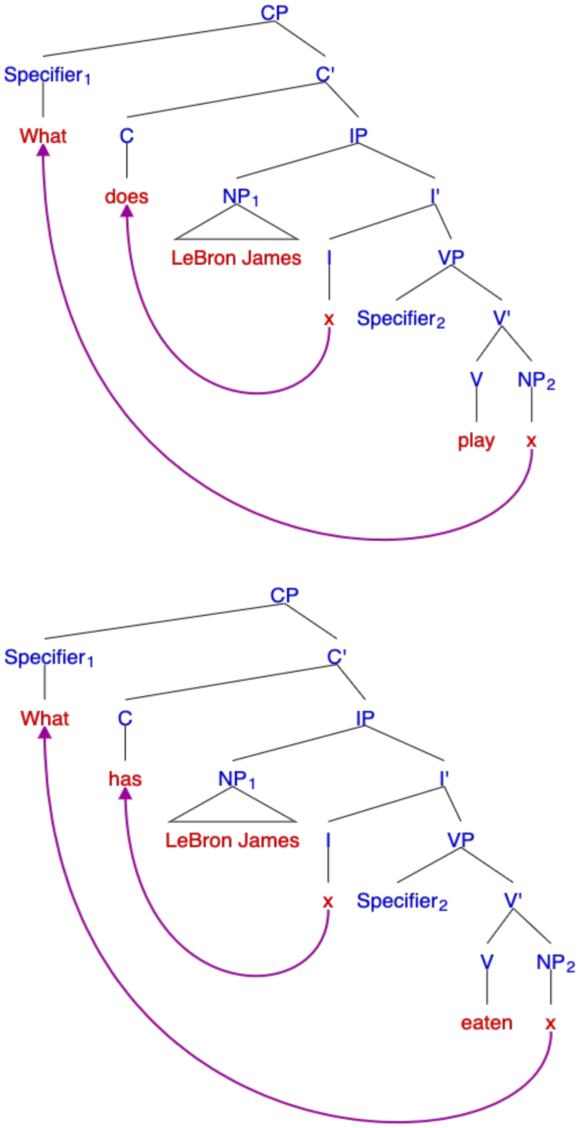


Figure 2 shows the formation of C and D. To form C, the wh-word (*what*) moves from its original position in IP (*LeBron James plays what*) to the specifier position of CP (specifier$_1$) (*What LeBron James plays*) and leaves a deletion trace. The rest of the process is similar to the formation of A. Based on the inflexional rules, the auxiliary DO (*does*, in this case) is generated automatically in head position and the inflexional suffix of the main verb *plays* (-s) disappears. It then moves to the head position of CP (*What does LeBron James play*) and leaves a deletion trace in the initial position. Similarly, to form D, the wh-word (*what*) moves from its original position in IP (*LeBron James has eaten what*) to the specifier$_1$ and leaves a deletion trace. The rest of the process is similar to the formation of B. Its auxiliary HAVE (*has*, in this case) then moves from the head position of IP to that of CP (*What has LeBron James eaten*) and leaves a deletion trace.

---

[4] Taco is a Mexican food, which is treated as an uncountable noun here.

Figure 2. Examples of formal representation of the formation of the non-subject wh-question including auxiliary DO and auxiliary HAVE



Based on the formation of these sample yes-no and non-subject wh-question, it can be found that all of them require the process of movement. Therefore, many generativist theories of child question acquisition maintains that movement is one of the principles contained in UG that children born with (Ambridge et al., 2006; Kania, 2016; C. F. Rowland, 2007; C. F. Rowland et al., 2005; Santelmann et al., 2002). According to this account, children quickly know English could allow movement operations (e.g., subject auxiliary inversion) in yes-no and non-subject wh-question under, albeit, limited exposure to English from their care givers (Rowland, 2007; Santelmann et al., 2002). This explains why children can produce adult-like questions from the very beginning of the language acquisition process (Rowland, 2007; Rowland et al., 2005), which is proved by the data (see Bellugi, 1965, 1971).

However, this account also has some problems. For example, there is a period when children are found to produce adult-like questions and non-adult-like questions including various errors at the same time (Ambridge et al., 2006). Theories appealing to early parameter setting are hard to explain this phenomenon, although they could successfully explain children's early adult-like question production (Rowland, 2007; Rowland et al., 2005; Santelmann et al., 2002). Furthermore, the question errors produced by children tend to show systematic patterns (Kania, 2016). For example, it was found in both corpus studies and experiments that children are more likely to make mistakes in questions including auxiliary DO (Hattori et al., 2003; Labov & Labov, 1978; Maratsos & Kuczaj, 1978; C. F. Rowland et al., 2005; Santelmann et al., 2002; Valian & Casey, 2003). A successful theory of child language acquisition (child question acquisition, in this case) should also explain these systematic errors in the early period of children's question production (Rowland, 2007). Nevertheless, although many generativists proposed different solutions (e.g., Bloom, 1990; De Villiers, 1991; Hyams, 1986; Radford, 1990, 1994; Santelmann et al., 2002; Stromwold, 1990), they still did not reach an agreement and constructed an integrated theory.

2.3. The generativist theory to explain children's question error patterns

There are mainly three influential generative accounts of questions error patterns (or error patterns from a broader view) in children's speech: the maturation theory (e.g., Babyonyshev et al., 2001; Borer & Wexler, 1992; Klima & Bellugi, 1966; Radford, 1990, 1994; Vainikka, 1993), the performance limitation theory (e.g., Bloom, 1990; Valian, 1991), and the integration theory (e.g., Santelmann et al., 2002; Stromwold, 1990).

The maturation theory posits that children produce non-adult-like sentences with errors in the multi-word speech stage because their brains have not matured enough to gain the full knowledge or full set of principles in UG, which means certain aspects of UG or certain principles have not been available to children yet (Babyonyshev et al., 2001; Borer & Wexler, 1987, 1992). Specifically, it is contended that movement or inversion is not available in children's early grammar (Klima & Bellugi, 1966; Radford, 1994); therefore, children is hard to utilise the grammatical knowledge of movement or inversion in their question formation, thus making various errors (Radford, 1994; Vainikka, 1993) (e.g., **what you are doing?*). However, there are also criticisms. For example, although the maturation theory predicts that the knowledge of functional categories in UG is not available to children at birth (Radford, 1994), it is found that children do rely on fictional categories when they begin to produce multi-word utterances (Lust, 1999). Moreover, Lust, (2006) argues that children master the tensed verb, determiner, and preposition in a period which is earlier than the maturation theory expects. Furthermore, if researchers suppose that the understanding of movement that allows inversion is a basic aspect of UG which is constantly accessible to children, then the alleged slow development of inversion in child grammar poses a challenge to the Strong Continuity Hypothesis of UG (see section 2.1 paragraph 3 for the detailed explanation) as a framework of language faculty of the children (Santelmann et al., 2002). The performance limitation theory claims that children's linguistic performance is limited by other immature cognitive abilities (e.g., working memory; attention), although they have access to all aspects of UG at first (Bloom, 1990; Valian, 1991). This idea is supported by many studies (e.g., Bloom, 1990; Hamburger & Crain, 1982). For example, some studies attribute children's difficulty in understanding relative clauses to the fact that they do not master relevant syntactic rules (e.g., Tavakolian, 1981). However, Hamburger and Crain (1982) reject this view and assert that task design confuses children. They redesigned the experiment and found that children understood relative clauses once given appropriate tasks. This indicates that researchers always underestimate children's grammatical competence due to their challenging experiment designs for children (Rowland, 2013). Nevertheless, the theory still fails to explain the auxiliary omission patterns (see Theakston et al., 2005) and the problem of lexical specificity (see Rowland, 2013 for detailed explanation).

The integration theory is believed to be more promising (Rowland, 2007). It also points out that children have access to all aspects of UG at birth, but they have to learn specific rules of inflexion (e.g., tense; number; person) in their mother language (e.g., English) and integrate them with the innate knowledge of movement to form questions (e.g., Santelmann et al., 2002; Stromwold, 1990). Santelmann et al. (2002) used an elicited imitation method to test the extent to which 2-to-5-year-old children master grammatical knowledge of inversion in English yes-no questions. The results showed that children could use the knowledge of inversion from the earliest tested age and did not change over time. They also showed the development of their knowledge of English inflexional rules. As Santelmann et al. (2002) predict, for example (see Figure 1 and section 2.2 for an detailed explanation of question formation), children will make more errors in questions including auxiliary DO because they have to learn "reconstruction of inflexion through do-support" (p. 814); in contrast, fewer errors would be produced in questions including auxiliary HAVE because children only need to utilise the innate knowledge of inversion to form such questions. Rowland (2007) further tested this theory through a corpus-based study. She examined children's percentage of errors in questions including auxiliary DO and modal auxiliaries and found that the percentage of errors of question including auxiliary DO was significantly higher than that with modal auxiliaries, which was consistent with Santelmann et al.'s (2002) study. However, she also noticed that yes-no questions account for a much more significant proportion than non-subject wh-question, which might affect the conclusion. Rowland then reanalysed the yes-no questions' and non-subject wh-question' percentage of errors independently and found that the percentage of errors of yes-no questions including auxiliary DO was also significantly higher than that with modal auxiliaries, in line with Santelmann et al.'s (2002) study. However, the results for the percentage of errors of non-subject wh-question were not as predicted. There was no significant difference between the mean percentage of errors of non-subject wh-question including auxiliary DO and that with modal auxiliaries (Rowland, 2007), although the researcher excluded the influences of the wh-word "why" (see Labov & Labov, 1978; Rowland et al., 2003; Rowland & Pine, 2000) and negative auxiliaries (see Bellugi, 1971; Guasti et al., 1995; Thornton & Houser, 2005).

Even though the generativist theory of child language acquisition successfully explains the question formation of children's utterances, it is still controversial whether it is applicable to explain various question error patterns in children's utterances, particularly in non-subject wh-question. Therefore, more empirical evidence from naturalistic data is needed to fill the research gap. This study aims to examine the generativist theory of child language acquisition to explain the question error patterns in children's utterances by examining 2-to-3-year-old children's percentage of errors of questions including auxiliary DO and auxiliary HAVE from their naturalistic speech. This study aims to address the following research questions.

RQ1: To what extent can the generativist theory explain children's overall error patterns in questions?

RQ2: To what extent can the generativist theory explain children's error patterns in yes-no questions and wh-questions respectively?

3. Methodology

This research adopts a cross-sectional corpus-based design to collect naturalistic data for quantitative analysis to compare percentage of errors in 2-to-3-year-old children's yes-no and non-subject wh-question including auxiliary DO and auxiliary HAVE.

3.1. Participants

The participants were chosen to collect required types of questions in a naturalistic setting. The participants were 12 British children (Anne, Aran, Becky, Carl, Dominic, Gail, Joel, John, Liz, Nicole, Ruth, Warren) from the Manchester corpus (Theakston et al., 2001) on CHILDES database (MacWhinney, 2000). All children are monolingual native English speakers from middle class families. All of them have typical language development paths and none of them suffer from cognitive problems or language disorders. Six of them are female and the rest of them are male. Table 1 [5] lists basic information of participants including the numbers of children, their names, their age ranges, and their MLU ranges. Their ages approximately range from 2 (1;08.22 - 2;00.25) to 3 (2;08.15 - 3;00.1). This age range is chosen because Santelmann et al.'s (2002) contends that the knowledge of inversion is available to children from the earliest testable age, i.e., the multi-word speech stage, which is around 2 years old (Lust, 2006). More information about children and can be found in the description page of the Manchester corpus and participants section in (Theakston et al., 2005).

Table 1. Basic information of participants

Number	Name	Age range	MLU* range
1	Anne	1;10.07 - 2;09.10	1.61 - 3.46
2	Aran	1;11.12 - 2;10.28	1.41 - 3.84
3	Becky	2;00.07 - 2;11.15	1.46 - 3.24
4	Carl**	1;08.22 - 2;08.15	2.17 - 3.93
5	Dominic**	1;10.24 - 2;10.16	1.20 - 2.85
6	Gail	1;11.27 - 2;11.12	1.76 - 3.42
7	Joel	1;11.01 - 2;10.11	1.33 - 3.32
8	John	1;11.15 - 2;10.24	2.22 - 2.93
9	Liz	1;11.09 - 2;10.18	1.35 - 4.12
10	Nicole**	2;00.25 - 3;00.10	1.06 - 3.26
11	Ruth**	1;11.15 - 2;11.21	1.41 - 3.35
12	Warren**	1;10.06 - 2;09.20	2.01 - 4.12

*MLU refers to the mean length of the utterance.

**Because Carl, Dominic, Nicole, Ruth, and Warren do not produce enough required types of questions, their data was excluded.

3.2. Transcription

Children's naturalistic utterances were transcribed orthographically. More details on the transcription can also be found in the description page of the Manchester corpus and transcription section in Theakston et al. (2005). The selected transcripts in the Manchester corpus are used for analysis.

3.3. Corpora

The corpora are built for required types of questions from the transcripts for coding and analysis. All yes-no questions including auxiliary HAVE and those including auxiliary DO from the transcripts were included in the corpora. They must contain the auxiliary, subject, and main verb. All non-subject wh-question including auxiliary HAVE and auxiliary DO from the transcripts were also incorporated in the corpora. They should have the wh-word, subject, auxiliary, and main verb. The questions including xxx marked in the transcripts (e.g., *where did he xxx?*) were excluded because it is hard to judge whether children are producing right (e.g., *where did he go?*) or non-adult-like questions (e.g., *where did he went?*). The subject non-subject Wh-question were not included in the corpora because children does not need inversion to produce such questions (e.g., *who did it?*), which is therefore irrelevant for the analysis. The non-inverted yes-no question were excluded (e.g., *you did it?*) because it is hard to judge whether it was a non-adult-like question without inversion or an intonation-only question produced on purpose.

3.4. Coding criteria

Coding criteria were adapted from Rowland et al. (2005) and Rowland (2007). All yes-no and non-subject wh-question including auxiliary DO or HAVE produced by the 12 children in the corpora in this study were coded by the researcher as follows.

3.4.1. Adult-like questions

In terms of adult-like questions, they were coded as (1) adult-like yes-no question including auxiliary DO, (2) adult-like yes-no questions including auxiliary HAVE, (3) adult-like non-subject wh-question including auxiliary DO, and (4) adult-like non-subject wh-question including auxiliary HAVE. All coded adult-like yes-no questions including auxiliary DO or HAVE should have the adult-like form and placement of the auxiliary, tense, agreement, case, main verb, and subject. All coded non-subject wh-question including auxiliary DO or HAVE should have the adult-like form and placement of the wh-word, tense, agreement, case, main verb, and subject. However, questions including omission and other minor grammatical errors were also coded as adult-like questions accordingly if they could show children's abilities to use inversion adult-likely in English. For example, *do you like play with dog?* is grammatically non-adult-like because *dog* should be plural or have the determiner, but it was also coded as adult-like yes-no question including auxiliary DO because *do you like play* already demonstrated children's good master of inversion in yes-no question formation.

---

[5] The information is checked and calculated by the researcher himself. However, Theakston et al. (2005) has a similar table showing more detailed information about children.

3.4.2. Non-adult-like questions

In terms of non-adult-like questions, they were coded as (1) non-adult-like yes-no question including auxiliary DO, (2) non-adult-like yes-no questions including auxiliary HAVE, (3) non-adult-like non-subject wh-question including auxiliary DO, and (4) non-adult-like non-subject wh-question including auxiliary HAVE. All coded non-adult-like yes-no questions should at least have the auxiliary, subject, and main verb. All coded non-adult-like non-subject wh-question should at least have the wh-word, auxiliary, and main verb, and they should have at least one grammatical problem in tense, agreement, case, subject omission, or inversion.

3.5. Data extraction

CLANc was used for the data extraction from the transcripts. The retrieval algorithm (combo +s"do+did+does+don't+didn't+doesn't" @ +t*CHI) was used for searching the transcripts for the utterances containing adult-like and non-adult-like yes-no and non-subject wh-question including auxiliary DO. Another retrieval algorithm (combo +s"have+has+had+haven't+hasn't+hadn't" @ +t*CHI) helped to search the transcripts for utterances including adult-like and non-adult-like yes-no and non-subject wh-question including auxiliary HAVE. All required adult-like and non-adult-like questions were then categorised by the researcher according to the coding criteria (see section 3.4).

3.6. Data analysis

SPSS 28 was the key software selected for the quantitative analysis. Descriptive data was calculated for the means and standard deviations of the numbers of children's different types of questions. Next, the means and standard deviations of the percentage of errors those questions were calculated accordingly (the percentage of errors = the number of non-adult-like questions / the total number of adult-like and non-adult-like questions). A paired sample t-test was then conducted to evaluate the influence of the auxiliary type (auxiliary DO or auxiliary HAVE) on the overall percentage of errors of children's questions. Another paired sample t-test were conducted to evaluate the effect of the question type (yes-no or non-subject wh-question) on the overall percentage of errors of children's questions. Next, a 2 x 2 within-subjects ANOVA was conducted with two factors. The first factor (the auxiliary type) had two levels (auxiliary DO or auxiliary HAVE), and the second factor (the question type) also had two levels (yes-no or non-subject wh-question). After that, a paired sample t-test was conducted to examine the impact of the auxiliary type on children's percentage of errors in yes-no questions. A paired sample t-test was conducted to examine the impact of the auxiliary type on children's percentage of errors in non-subject wh-question. A paired sample t-test was conducted to examine the impact of the question type on children's percentage of errors in questions including auxiliary DO. Another paired sample t-test was conducted to examine the impact of the question type on children's percentage of errors in questions including auxiliary HAVE.

4. Results

4.1. Error patterns in questions including auxiliary DO and auxiliary HAVE

Table 2 shows the mean and standard deviation of the number of adult-like and non-adult-like questions including auxiliary DO and auxiliary HAVE and the percentage of errors of questions including auxiliary DO and auxiliary HAVE. It can be observed that the mean percentage of errors with questions of auxiliary DO (5.78%) is higher than that with auxiliary HAVE (2.03%). However, due to large standard deviation, a paired sample t-test was conducted to examine the impact of auxiliary type (auxiliary DO or auxiliary HAVE) on children's percentage of errors in questions. The results show that there is no significant difference between children's percentage of errors of questions including auxiliary DO and those including auxiliary HAVE (t (6) = 1.565, p = .169, two-tailed, 95% confidence interval level [-.02116, .09620]). The Cohen's d (.063 < .2) indicates a very small effect size (Pallant, 2020).

Table 2. The mean (M) and standard deviation (SD) of the number of adult-like and non-adult-like questions including auxiliary DO and auxiliary HAVE and the corresponding percentages of errors

	Questions including auxiliary DO			Questions including auxiliary HAVE		
	N_a[6]	N_n[7]	Percentage of errors (%)	N_a	N_n	Percentage of errors (%)
Mean	91.14	8.86	5.78	17.86	.86	2.03
(SD)	(110.64)	(18.25)	(6.76)	(15.78)	(1.86)	(3.62)

Furthermore, Rowland (2007) claims that the question type (yes-no or non-subject wh-question) can impact children's percentage of errors in questions, and it is better to analyse children's percentage of errors of yes-no and non-subject wh-question separately. Therefore, a paired sample t-test was conducted to evaluate the influence of the question type (yes-no or non-subject wh-question) on children's percentage of errors. The data shows that children's percentage of errors differ significantly in yes-no questions (M = .0348, SD = .01925) and non-subject wh-question (M = .1155, SD = .03051, t (6) = -3.327, p = .016 < .05, two-tailed) with a 95% confidence interval (CI) level ranging from -.14016 to -.02135 (see Table 3). However, the effect size is very small (Cohen'd = .064 < .2) (Pallant, 2020). It indicates that the question type could have an impact on children's percentage of errors in questions.

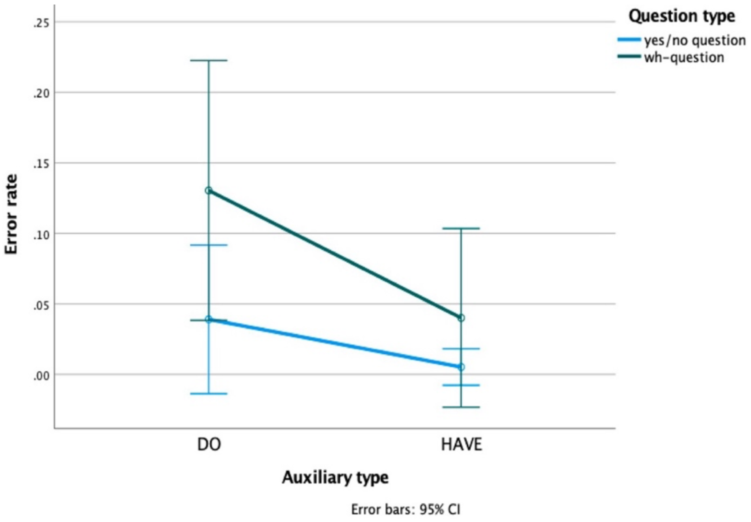
---

[6] This symbol refers to the number of adult-like questions in this paper.

[7] This symbol refers to the number of non-adult-like questions in this paper.

To further confirm the impact of the question type and the auxiliary type on the percentage of errors and examine their interaction effect, a 2 x 2 within-subjects ANOVA was conducted with two factors (see Figure 3). The first factor (the auxiliary type) had two levels (auxiliary DO or auxiliary HAVE), and the second factor (the question type) also had two levels (yes-no or non-subject wh-question). The results show that the main effect of the auxiliary type is significant (F (1, 6) = 6.298, p = .046 < .05). The size effect is very large (η^2 = .509 > .138) (ibid.). The main effect of the question type is also significant (F (1, 6) = 10.913, p = .016 < .05). The size effect is very large (η^2 = .651 > .138) (Pallant, 2020). However, there is no interaction effect between the auxiliary type and the question type (F (1, 6) = 2.039, p = .203 > .05). The size effect is very large (η^2 = .273 > .138) (Pallant, 2020). Figure 3 shows children's percentage of errors in yes-no and non-subject wh-question including auxiliary DO and auxiliary HAVE respectively. The two lines shows the similar tendency, implying that there is less likely to have an interaction effect between two factors (Harrison *et al.*, 2022, p. 280), which is consistent with the statistical data. Due to the main effect of the question type, children's percentage of errors of questions including auxiliary DO and auxiliary HAVE were analysed separately below according to the question type.

Figure 3. Children's percentage of errors (with error bars) in yes-no and non-subject wh-question including auxiliary DO and auxiliary HAVE



4.2. Error patterns of yes-no questions and non-subject wh-questions including auxiliary DO and auxiliary HAVE

Table 3 presents the information about the mean and standard deviation of the number of adult-like and non-adult-like yes-no questions including auxiliary DO and auxiliary HAVE and the percentage of errors of yes-no questions including auxiliary DO and auxiliary HAVE. It can be found that the mean percentage of errors of yes-no questions including auxiliary DO (3.91%) is larger than that of yes-no questions including auxiliary HAVE (0.53%). A paired sample t-test was conducted to examine the impact of the auxiliary type on children's percentage of errors in yes-no questions. The data shows that there is no significant difference in children's percentage of errors in yes-no questions including auxiliary DO (M = .0391, SD = .05693) and those including auxiliary HAVE (M = .0053, SD = .01400, t (6) = 1.685, p = .143 > .05, two-tailed) with a 95% confidence interval level ranging from -.01527 to -.08281 (see Table 5). However, the effect size is very small (Cohen'd = .053 < .2).

Table 3. The mean and standard deviation of the number of adult-like and non-adult-like yes-no questions including auxiliary DO and auxiliary HAVE and the corresponding percentages of errors

	Yes-no questions including auxiliary DO			Yes-no questions including auxiliary HAVE		
	N_a	N_n	Percentage of errors (%)	N_a	N_n	Percentage of errors (%)
Mean	73.00	4.43 (10.42)	3.91	12.00	.14	.53
(SD)	(95.31)		(5.69)	(8.64)	(.38)	(1.40)

Table 4 presents the information about the mean and standard deviation of the number of adult-like and non-adult-like non-subject wh-question including auxiliary DO and auxiliary HAVE and the percentage of errors of non-subject wh-question including auxiliary DO and auxiliary HAVE. It can be found that the mean percentage of errors of non-subject wh-question including auxiliary DO (13.05%) is larger than that of non-subject wh-question including auxiliary HAVE (4.01%). A paired sample t-test was conducted to examine the impact of the auxiliary type on children's percentage of errors in non-subject wh-question. The data shows that children's percentage of errors in non-subject wh-question including auxiliary DO (M = .1305, SD = .0996) is significantly larger than those including auxiliary HAVE (M = .0401, SD = .026, *t* (6) = 1.685, *p* = .033 < .05, one-tailed) with a 95% confidence interval level ranging from -.00772 to -.1884. However, the effect size is very small (Cohen'd = .106 < .2) (Pallant, 2020).

Table 4. The mean and standard deviation of the number of adult-like and non-adult-like yes-no questions including auxiliary DO and auxiliary HAVE and the corresponding percentages of errors

	Non-subject wh-question including auxiliary DO			Non-subject wh-question including auxiliary HAVE		
	N_a	N_n	Percentage of errors (%)	N_a	N_n	Percentage of errors (%)
Mean	18.14	4.43	13.05	5.86	.71	4.01
(SD)	(17.112)	(7.913)	(.0996)	(8.688)	(1.496)	(.026)

4.3. The impact of the question type on children's percentage of errors in questions including auxiliary DO and HAVE

To further examine the impact of the question type on children's percentage of errors in questions including auxiliary DO and HAVE, a paired sample t-test was conducted to examine the impact of the question type on children's percentage of errors in questions including auxiliary DO (see Table 5). The data shows that children's percentage of errors in yes-no questions including auxiliary DO (M = .0391, SD = .05693) is significantly lower than non-subject wh-question including auxiliary DO (M = .1305, SD = .09957, *t* (6) = -2.907, *p* = .027 < .05, two-tailed) with a 95% confidence interval level ranging from -.16833 to -.01446 (see Table 6). However, the effect size is very small (Cohen'd = .083 < .2) (Pallant, 2020).

Table 5. The result of the paired sample t-test to compare percentage of errors in yes-no and non-subject wh-question including auxiliary DO (two-tailed)

		M	SD	*t*	Sig.	95% CI	
						Lower	Upper
Pair	Auxiliary DO (yes-no)	.0391	.05693	-2.907	.027	-.16833	-.01446
	Auxiliary DO (wh)	.1305	.09957				

Another paired sample t-test was conducted to examine the impact of the question type on children's percentage of errors in questions including auxiliary HAVE (see Table 6). The data shows that There is no significant difference between children's percentage of errors in yes-no questions including auxiliary HAVE (M = .0053, SD = .01400) and non-subject wh-question including auxiliary HAVE (M = .0401, SD = .06852, *t* (6) = -1.518, *p* = .180 > .05, two-tailed) with a 95% confidence interval level ranging from -.09095 to .02131 (see Table 9). However, the effect size is very small (Cohen'd = .06069 < .2) (Pallant, 2020).

Table 6. The result of the paired sample t-test to compare percentage of errors in yes-no and non-subject wh-question including auxiliary HAVE (two-tailed)

		M	SD	*t*	Sig.	95% CI	
						Lower	Upper
Pair	Auxiliary HAVE (yes-no)	.0053	.0140	-1.518	.180	-.09095	.02131
	Auxiliary HAVE (wh)	.0401	.0685				

5. Discussion

5.1. To what extent can the generativist theory explain the error patterns in questions?

It is found that the mean percentage of errors of questions including auxiliary DO is higher than that with auxiliary HAVE, although there is no significant difference between children's percentage of errors of questions including auxiliary DO and those including auxiliary HAVE with a very small size effect. This is partially consistent with Santelmann et al.'s (2002), who proposes the integration theory, experiments that indicate that the percentage of errors in questions including auxiliary DO are more likely to be higher. It is also in line with other studies contending that children are more likely to make errors in questions including auxiliary DO (Hattori et al., 2003; Labov & Labov, 1978; Maratsos & Kuczaj, 1978; C. F. Rowland et al., 2005; Valian & Casey, 2003). Moreover, according to the integration theory, children should produce less errors in questions including the modal auxiliaries and auxiliary HAVE because questions including then do not require the integration of infection system to be produced. Therefore, the finding also echoes the finding in Rowland's (2007) corpus study that percentage of errors in questions including modal auxiliaries is lower than those including auxiliary DO. It is also found that the main effect of the auxiliary type is significant with a large size effect. This is against Rowland's (2007) finding that the main effect of the auxiliary type is not significant. However, this finding is consistent with the finding that auxiliary DO attracts a higher percentage of errors. In conclusion, the findings indicate that the generativist theory could successfully explain children's general percentage of errors in questions.

Moreover, it is found that children's percentage of errors differ significantly in yes-no questions and non-subject wh-question and the main effect of the question type is significant with a large size effect. Santelmann et al. (2002) did not notice the impact of the question type. However, Rowland (2007) found and emphasised the impact of question type: the main effect of the question type is significant, which is in line with the findings in the current study. Furthermore, it is also found that children's mean percentage of errors in yes-no questions including auxiliary DO is significantly lower than non-subject wh-question including auxiliary DO and the mean percentage of errors in non-subject wh-question including auxiliary DO is also lower than non-subject wh-question including auxiliary DO, although there is no significant difference between children's percentage of errors in yes-no questions including auxiliary HAVE and non-subject wh-question including auxiliary HAVE with a very small size effect. This further confirm the impact of the question type. Moreover, this study also finds that there is no interaction effect between the auxiliary type and the question type with a large size effect, which is contrary to Rowland's (2007) finding that the interaction effect is highly significant. It is hard to explore the implication based on the interaction effect, and it is therefore inconclusive here. In conclusion, the findings indicate that the generativist explanation of children's question error pattens should be constructed based on different question type (yes-no or wh-question).

5.2. To what extent can the generativist theory explain the error patterns in yes-no questions?

It is found that the mean percentage of errors of yes-no questions including auxiliary DO is larger than that of yes-no questions including auxiliary HAVE, although there is no significant difference in children's percentage of errors in yes-no questions including auxiliary DO and those including auxiliary HAVE and the effect size is very small. This finding is consistent with the prediction from Santelmann et al. (2002). Together with Rowland's finding that children's percentage of errors in questions including modal auxiliaries is higher than those including auxiliary DO, the finding in the current study could support the integration theory from generativists (e.g., Santelmann et al., 2002; Stromwold, 1990). This indicates that the generativist theory could also successfully explain the error patterns in yes-no questions.

This study also finds that the mean percentage of errors of non-subject wh-question including auxiliary DO is larger than that of non-subject wh-question including auxiliary HAVE. Moreover, children's percentage of errors in non-subject wh-question including auxiliary DO is significantly larger than those including auxiliary HAVE. This finding echoes the finding in Rowland et al.'s (2005) study that auxiliary HAVE attracts lower percentage of errors than auxiliary DO in non-subject wh-question formation. However, in terms of the generativist account, this finding is against Rowland's (2007) study which contents that children produce more errors in non-subject wh-question including modal auxiliaries than those including auxiliary DO because based on the integration theory, wh-question including auxiliary DO should attract higher percentage of errors due to integration with inflexional system in formation. Moreover, although many studies challenge the generativist theory to explain children's error pattern in wh-question and proposed a constructivist solution (Rowland, 2007; Rowland et al., 2005; Rowland & Pine, 2000), the finding in the current study supports the integration theory which proposed by the generativists. This indicates that the generativist theory could also successfully explain the error patterns in non-subject wh-question.

6. CONCLUSION

In conclusion, this study aims to explore whether the generativist account, specifically the integration theory could explain children's percentage of errors in question in general whether it is also applicable to yes-no and wh-question. The current study adopts a corpus-based method to compare 2-to-3-year-old children's percentage of errors in questions (and yes-no and wh-question separately) with auxiliary DO and auxiliary HAVE. The results show that (1) the mean percentage of errors of questions including auxiliary DO is higher than that with auxiliary HAVE, although there is no significant difference between children's percentage of errors of questions including auxiliary DO and those including auxiliary HAVE with a very small size effect; (2) The mean percentage of errors of yes-no questions including auxiliary DO is larger than that of yes-no questions including auxiliary HAVE, although there is no significant difference in children's percentage of errors in yes-no questions including auxiliary DO and those including auxiliary HAVE and the effect size is very small; (3) The mean percentage of errors of non-subject wh-question including auxiliary DO is larger than that of non-subject wh-question including auxiliary HAVE and children's percentage of errors in non-subject wh-question including auxiliary DO is significantly larger than those including auxiliary HAVE. Therefore, the current study concludes that the generativist theory could successfully explain children's overall percentage of errors in questions and percentage of errors in yes-no and non-subject wh-question. This study provides empirical evidence to support the generativist theory of child question acquisition and theory of child language acquisition in a broader view. Moreover, this study also finds that children's percentage of errors differ significantly in yes-no questions and non-subject wh-question and the main effect of the question type is significant with a large size effect, which indicates that the generativist explanation of children's question error pattens should be constructed based on different question type (yes-no or non-subject wh-question). This finding provides insights for improving and refining the generativist theory of child question acquisition.

Admittedly, this study does have some problems. Firstly, the sample size is small (only seven participants) due to exclusion of the participants who do not produce enough required types of data, which cause statistical insignificance and very small effect size in many statistical tests in this study. However, the small sample size is common in cross-sectional corpus-based study to explore child question acquisition (e.g., 12 participants in Rowland et al., 2003; 13 participants in Rowland et al., 2005; 10 participants in Rowland, 2007) because it is time-consuming and effortful to collect and analyse data (Rowland et al., 2008). Therefore, this study uses the mean as an important indicator of children's tendency to produce errors in different types of questions including the help of statistical techniques as supplements. However, later research should still consider carefully about the problem of sample size. Moreover, this study also finds that there is no interaction effect between the auxiliary type and the question type with a large size effect, which is contrary to Rowland's (2007) finding that the interaction effect is highly significant. It is hard to explore the implication based on the interaction effect, and it is therefore inconclusive here. Further research could follow this idea and design experiments accordingly to figure out the implications of the interaction effect between the question type and the auxiliary effect in children's questions.

Acknowledgement

Many thanks to my supervisor, Dr Ursula Kania, for her kind suggestions. Thanks for my parents' emotional and financial support for my academic career.

Yiran Du is currently an undergraduate in University of Liverpool and will study in University of Oxford (MPhil in Linguistics, Philology and Phonetics). His research interests are psycholinguistics, semantics, and language acquisition.
E-mail: hsydu14@liverpool.ac.uk

References

Ambridge, B., Rowland, C. F., Theakston, A. L., & Tomasello, M. (2006). Comparing different accounts of inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?' *Journal of Child Language*, *33*(3), 519–557. https://doi.org/10.1017/S0305000906007513

Babyonyshev, M., Ganger, J., Pesetsky, D., & Wexler, K. (2001). The Maturation of Grammatical Principles: Evidence from Russian Unaccusatives. *Linguistic Inquiry*, *32*(1), 1–44. https://doi.org/10.1162/002438901554577

Bellugi, U. (1965). The Development of Interrogative Structures in Children's Speech. In K. Riegel (Ed.), *The development of language functions* (pp. 103–138). University of Michigan.

Bellugi, U. (1971). Simplification in children's language. In R. Huxley & E. Ingram (Eds.), *Language acquisition: Models and methods* (pp. 95–119). Academic Press.

Bloom, P. (1990). Subjectless Sentences in Child Language. *Linguistic Inquiry*, *21*(4), 491–504.

Borer, H., & Wexler, K. (1987). The Maturation of Syntax. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (Vol. 4, pp. 123–172). Springer Netherlands. https://doi.org/10.1007/978-94-009-3727-7_6

Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory*, *10*(2), 147–189. https://doi.org/10.1007/BF00133811

Chomsky, N. (1957). *Syntactic Structures*. Mouton.

Chomsky, N. (1959). *Language*, *35*(1), 26–58. JSTOR. https://doi.org/10.2307/411334

Chomsky, N. (1964). *Current Issues in Linguistic Theory*. Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT Press.

Chomsky, N. (1980). *Rules and representations*. Columbia University Press.

Chomsky, N. (1981). *Lectures on Government and Binding: The Managua lectures*.

Chomsky, N. (1986). *Knowledge of language*. Praeger.

Chomsky, N. (1995). *The minimalist grogram*. MIT Press.

Clahsen, H. (1990). Constraints on Parameter Setting: A Grammatical Analysis of Some Acquisition in Stages in German Child Language. *Language Acquisition*, *1*(4), 361–391. https://doi.org/10.1207/s15327817la0104_3

De Villiers, J. (1991). Why Questions? *University of Massachusetts Occasional Papers in Linguistics*, *17*(1), 155–173.

Erreich, A. (1984). Learning how to ask: Patterns of inversion in *yes – no* and *wh* -questions. *Journal of Child Language*, *11*(3), 579–592. https://doi.org/10.1017/S0305000900005961

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology. A Bradford book*. MIT Press.

Fodor, J. D., & Sakas, W. G. (2004). *Evaluating models of parameter setting*. *1*, 1–27.

Garnham, A. (2013). Language, the Mind, and the Brain. In K. Allan (Ed.), *The Oxford Handbook of the History of Linguistics* (1st ed., pp. 674–690). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199585847.013.0031

Guasti, M. T., Thornton, R., & Wexler, K. (1995). Children's negative questions: The case of English. *Proceedings of the Boston University Conference on Language Development*, *19*, 228–

239.

Haan, G. de. (1987). A theory-bound approach to the acquisition of verb placement in Dutch. In G. de. Haan & W. Zonneveld (Eds.), *Formal parameters of generative grammar* (pp. 15–30). University of Utrecht.

Hamburger, H., & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Ed.), *Language development: Syntax and semantics* (pp. 245–274). Erlbaum.

Hattori, R., Alejn, B., Micciulla, L., & Smith, C. E. (2003). Why do children say did you went?: The role of do-support. *Supplement to the Proceedings of the 28th Boston University Conference on Language Development*.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, *298*(5598), 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Reidel.

Ingram, D., & Tyack, D. (1979). Inversion of subject NP and aux in children's questions. *Journal of Psycholinguistic Research*, *8*(4), 333–341. https://doi.org/10.1007/BF01067137

Jordens, P. (1990). The acquisition of verb placement in Dutch and German. *Ling*, *28*(6), 1407–1448. https://doi.org/10.1515/ling.1990.28.6.1407

Kania, U. (2016). *The acquisition and use of yes-no questions in English: A corpus-study from a usage-based perspective*. Narr Francke Attempto.

Klee, T. (1985). Role of Inversion in Children's Question Development. *Journal of Speech, Language, and Hearing Research*, *28*(2), 225–232. https://doi.org/10.1044/jshr.2802.225

Klima, E. S., & Bellugi, U. (1966). Syntactic Regulation in the Speech of Children. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistic Papers* (pp. 183–203). Edingurgh University Press.

Kuczaj, S. A. (1976). Arguments against Hurford's 'Aux Copying Rule'. *Journal of Child Language*, *3*(3), 423–427. https://doi.org/10.1017/S0305000900007297

Labov, W., & Labov, T. (1978). Learning the Syntax of Questions. In R. N. Campbell & P. T. Smith (Eds.), *Recent Advances in the Psychology of Language* (pp. 1–44). Springer US. https://doi.org/10.1007/978-1-4684-2532-1_1

Lust, B. C. (1999). Universal grammar: The strong continuity hypothesis in first language acquisition. In *Handbook of child language acquisition.* (pp. 111–155). Academic Press.

Lust, B. C. (2006). *Child Language: Acquisition and Growth*. Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum.

Maratsos, M., & Kuczaj, S. A. (1978). Against the transformationalist account: A simpler analysis of auxiliary overmarkings. *Journal of Child Language*, *5*(2), 337–345. https://doi.org/10.1017/S0305000900007510

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, *7*(3), 141–144. https://doi.org/10.1016/S1364-6613(03)00029-9

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, *65*(3), 151–166. https://doi.org/10.1037/h0048495

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th edition). Open University Press.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.

Pinker, S. (1984). *Language Learnability and Language Development*. Harvard University Press.

Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, *18*(1–2). https://doi.org/10.1515/tlir.19.1-2.9

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Blackwell.

Radford, A. (1994). The syntax of questions in child English. *Journal of Child Language*, *21*(1), 211–236. https://doi.org/10.1017/S0305000900008722

Rowland, C. (2013). *Understanding Child Language Acquisition*. Routledge. https://doi.org/10.4324/9780203776025

Rowland, C. F. (2007). Explaining errors in children's questions. *Cognition*, *104*(1), 106–134. https://doi.org/10.1016/j.cognition.2006.05.011

Rowland, C. F., Fletcher, S. L., & Freudenthal, D. (2008). How big is big enough? Assessing the reliability of data from naturalistic samples. In H. Behrens (Ed.), *Corpora in Language Acquisition Research: History, methods, perspectives*. John Benjamins Publishing Company.

Rowland, C. F., & Pine, J. M. (2000). Subject–auxiliary inversion errors and wh-question acquisition: 'What children do know?' *Journal of Child Language*, *27*(1), 157–181. https://doi.org/10.1017/S0305000999004055

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language*, *30*(3), 609–635. https://doi.org/10.1017/S0305000903005695

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2005). The Incidence of Error in Young Children's *Wh* -Questions. *Journal of Speech, Language, and Hearing Research*, *48*(2), 384–404. https://doi.org/10.1044/1092-

4388(2005/027)

Santelmann, L., Berk, S., Austin, J., Somashekar, S., & Lust, B. (2002). Continuity and development in the acquisition of inversion in yes/no questions: Dissociating movement and inflection. *Journal of Child Language*, *29*(4), 813–842. https://doi.org/10.1017/S0305000902005299

Skinner, B. F. (1957). *Verbal Behavior*. Appleton-Century-Crafts.

Stromwold, K. (1990). *Learnability and the acquisition of auxiliaries*. MIT Working Papers in Linguistics.

Tavakolian, S. (1981). The conjoined clause analysis of relative clauses. In *Language acquisition and linguistic theory* (pp. 167–187). MIT Press.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*(1), 127–152. https://doi.org/10.1017/S0305000900004608

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2005). The acquisition of auxiliary syntax: BE and HAVE. *Cognitive Linguistics*, *16*(1). https://doi.org/10.1515/cogl.2005.16.1.247

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan: Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, *21*(3), 217–228. https://doi.org/10.1111/j.1365-2729.2005.00129.x

Tomasello, M. (2005). Beyond formalities: The case of language acquisition. *The Linguistic Review*, *22*(2–4). https://doi.org/10.1515/tlir.2005.22.2-4.183

Traxler, M. J. (2016). Psycholinguistics: Language and cognition. In K. Allan (Ed.), *The Routledge Handbook of Linguistics* (pp. 281–295). Routledge.

Vainikka, A. (1993). Case in the Development of English Syntax. *Language Acquisition*, *3*(3), 257–325. https://doi.org/10.1207/s15327817la0303_3

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1–2), 21–81. https://doi.org/10.1016/0010-0277(91)90046-7

Valian, V., & Casey, L. (2003). Young children's acquisition of *wh* -questions: The role of structured input. *Journal of Child Language*, *30*(1), 117–143. https://doi.org/10.1017/S0305000902005457

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*(1), 1–14. https://doi.org/10.1037/h0069608