

Systematic Review

Endangered Tanka Language of the Maritime Communities Across Southeast China: Convergence and Loss

Yanmei Dai  and Cong Wang * 

Department of Chinese Language Studies, The Education University of Hong Kong, New Territories 999077, Hong Kong

* Correspondence: wangc@eduhk.hk

Abstract

Amidst global concerns for linguistic diversity, this systematic review synthesizes six decades (1965–2025) of research on Tanka, a critically endangered language spoken by the boat people along Southeast China. Analyzing 42 studies identified through the PRISMA framework, the review reveals significant sociolinguistic and epistemological imbalances. Research output disproportionately focuses on phonetics and phonology (50%), while neglecting grammar, lexicon, and sociolinguistic vitality. Linguistically, Tanka demonstrates substantial contact-induced convergence with Cantonese or Pinghua within multilingual ecologies; nevertheless, it retains distinctive phonological shifts, a unique maritime lexicon, and grammatical innovations, reflecting both regional alignment and endogenous community practices. Its heterogeneous genetic affiliation highlights local sociohistorical contact dynamics. Rapid intergenerational language shift is documented across communities, driven by intersecting pressures, including state-led urbanization, Mandarin-centric education policies, demographic shifts, occupational change, and enduring social stigmatization. Therefore, community attitudes often prioritize socio-economic mobility through dominant languages over heritage maintenance. Persistent gaps include limited syntactic and discourse analysis, minimal use of quantitative and computational methods (e.g., AI-assisted documentation), insufficient geographic coverage, and a lack of longitudinal shift studies. The field thus urgently requires enhanced international engagement via English publications and a decisive shift towards collaborative, community-centered revitalization frameworks that address power asymmetries and harness cultural resilience.

Keywords: Tanka language; language endangerment; linguistic convergence; maritime communities; Southeast China



Academic Editor: Anvita Abbi

Received: 9 December 2025

Revised: 20 February 2026

Accepted: 25 February 2026

Published: 5 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The boat people, also known as “Tanka people” or “water gypsies”, are communities who have lived aboard boats for generations and were historically prohibited from settling on land (McCoy, 1965; Kiu, 1984; Wang et al., 2025). Sustained by fishing, aquatic trade, and maritime livelihoods, they maintained fluid boundaries between land-based and waterborne life across Southeast China, particularly in Guangdong, Guangxi, Hainan, Hong Kong, and Macau (Bai, 2007a; McCoy, 1965; Fung, 2023; Pu, 2019; Wang et al., 2025).

The Tanka language, spoken by these boat-dwelling communities, embodies a rich maritime epistemology, encoding specialized knowledge of fishing techniques, boat architecture, weather systems, and ecological practices (Bai, 2007b; Cheng, 2020; Fung, 2015;

Wang et al., 2025). Despite its profound cultural significance, the language is now critically endangered due to urbanization, assimilation policies, and intergenerational language shift (Chan, 2012; G. F. Huang & Xu, 2018; Wang et al., 2025). Younger generations increasingly adopt Cantonese or Mandarin, while Tanka usage has become restricted to elder speakers and ritual domains (Cheng, 2020; Wang et al., 2025). With limited revitalization initiatives and uneven documentation, Tanka faces an acute risk of extinction, mirroring global patterns of minority language decline (Fishman, 1991; Bradley & Bradley, 2002).

Although individual studies have described phonetic and lexical similarities between Tanka and regional Cantonese or Pinghua varieties (Chan, 2012; Z. X. Huang, 2015; Zhuang, 2009), research to date remains fragmented, regionally uneven, and methodologically narrow. Most works focus on isolated features, such as phonology or cultural lexicon, without integrating findings across linguistic domains or geographic regions. Consequently, there is still no systematic review that situates Tanka within the broader Sinitic typological network or examines its contact-induced evolution in the multilingual ecology of Southeast China. This absence of systematic integration has limited theoretical engagement and obscured cross-regional patterns essential for understanding Tanka's genealogical affiliation, structural variability, and sociolinguistic transformation. To address these gaps and establish a unified framework for analysis, this review is guided by the following research questions (RQs):

RQ1: What are the major trends, thematic focuses, and methodological approaches in existing studies on the Tanka language?

RQ2: What insights have recent studies provided regarding the phonetic, lexical, and grammatical features of the Tanka language, its genetic affiliation, and its interactions with other Sinitic languages?

RQ3: How have studies documented and interpreted the roles of identity negotiation, language attitudes, and sociolinguistic factors in shaping language use and intergenerational transmission within Tanka-speaking communities, and what implications do these findings have for language shift and preservation?

RQ4: What are the main limitations of current research on the Tanka language, and what future directions should studies pursue to address these gaps?

2. Methodology

2.1. Systematic Review Protocol

This study adopts a systematic review methodology consistent with PRISMA guidelines (Moher et al., 2009; Page et al., 2021) to ensure a transparent and reproducible process for identifying, selecting, and synthesizing research on the Tanka language. To further reinforce methodological clarity and analytical coherence, the review is organized according to the PICOS model, which specifies the Population studied, the Intervention or focus of analysis, the Comparator or reference framework, the Outcomes examined, and the Study Designs included.

2.2. Data Sources and Search Strategy

A two-phase search strategy combined structured database queries with citation tracking to capture both English- and Chinese-language scholarship. Searches were conducted in Web of Science, Scopus, and CNKI, selected for their disciplinary breadth and bilingual coverage. A Boolean keyword strategy was used to identify relevant studies related to the Tanka language and associated cultural-linguistic phenomena. The search query was as follows:

TITLE-ABS-KEY ("Tanka language" OR "Tanka phonetics*" OR "Tanka vocabulary*" OR "Tanka grammar*" "boat people" OR "Danjia"¹ OR "Affiliation" OR "language attitude*" OR "language use") AND PUBYEAR > 1965 AND (LIMIT-TO (SUBJAREA, "SOC") OR LIMIT-TO (SUBJAREA, "ARTS") AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "ch") OR LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "bk") AND (LIMIT-TO (LANGUAGE, "English" OR "Chinese").

The search was completed on 1 July 2025. All retrieved records were imported into Zotero for deduplication and preliminary screening based on titles and abstracts.

Both forward and backward citation tracking were conducted to ensure the comprehensiveness of the literature search. Forward citation tracking was carried out using the "Cited by" functions in Web of Science, Scopus, CNKI, and Google Scholar to identify recent studies that referenced key foundational works, such as McCoy (1965) and Kiu (1984). Backward citation tracking, implemented through reference snowballing, involved reviewing the bibliographies of seminal publications, such as Fung (2015), Yao (2022), and Wang et al. (2025), to locate earlier or thematically related studies. This approach minimized omissions due to terminological or indexing inconsistencies.

2.3. Inclusion and Exclusion Criteria

All retrieved studies were independently screened by two authors based on pre-defined inclusion and exclusion criteria. Any discrepancies in judgment were resolved through discussion until consensus was reached, ensuring methodological rigor and consistency.

(1) Inclusion criteria

- (1) Subject areas: studies focusing primarily on the Tanka language, including its phonetics, phonology, morphology, syntax, vocabulary, and classification within the Sinitic language family, and analyses that contextualize linguistic features within sociocultural or ethnolinguistic frameworks, language identity, attitudes, and language use.
- (2) Publication year: between 1965 and 2025.
- (3) Document types: Published works, including peer-reviewed journal articles, academic monographs, conference papers, and theses.
- (4) Languages: Chinese and English.

(2) Exclusion criteria:

- (1) Works solely discussing other Sinitic or maritime languages (e.g., Hoklo) without comparative or specific reference to Tanka.
- (2) Anthropological, cultural or genetic studies of Tanka people that lack explicit linguistic analysis.
- (3) Studies lacking sufficient methodological or linguistic detail.
- (4) Non-academic or unpublished materials, such as newspapers, blog posts, folklore retellings without linguistic annotation, unpublished conference presentations, and social media discussions.

A PRISMA 2020 flow diagram (see Figure 1) illustrates the study selection process, detailing the number of records identified, screened, assessed for eligibility, and ultimately included in the final analysis.

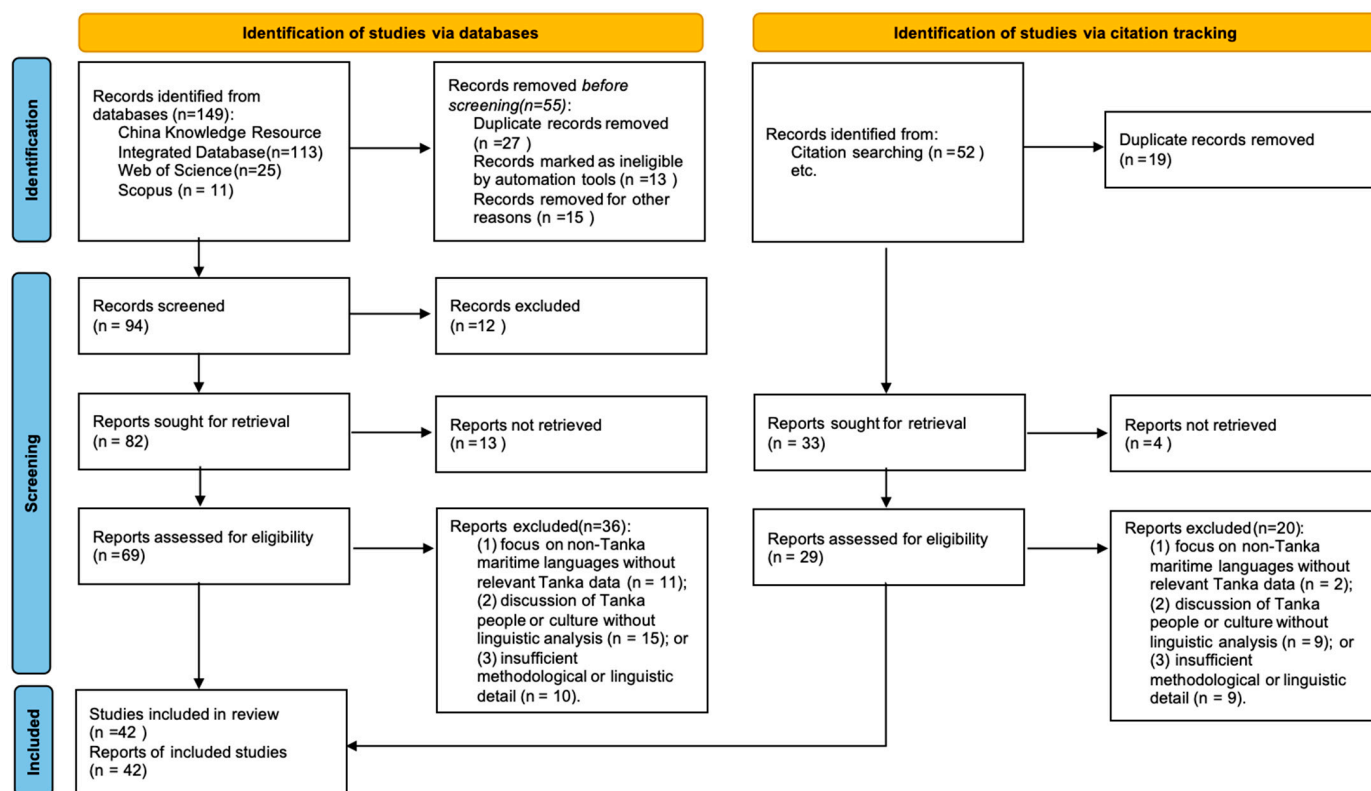


Figure 1. PRISMA 2020 flow diagram for the systematic review of Tanka language studies.

2.4. Quality Assessment

To ensure methodological rigor and analytical validity, all sources were appraised using criteria adapted from the Critical Appraisal Skills Programme (CASP, 2018) and informed by best practices in linguistic and sociocultural synthesis (Tong et al., 2007; Tricco et al., 2018). Each study was evaluated across five dimensions: (1) clarity of research objectives and their alignment with design and sociolinguistic context; (2) transparency of data collection and sampling procedures; (3) credibility of data and reliability of linguistic transcriptions, including IPA and tone notation; (4) analytical coherence and theoretical grounding, with attention to reflexivity and methodological limitations; and (5) peer-review status and institutional credibility. Only published, peer-reviewed articles, monographs, conference papers, and theses, submitted to accredited universities, supervised by recognized linguists and archived in accessible academic repositories, were included. Studies that did not meet these standards were excluded. This systematic yet context-sensitive appraisal ensured both academic integrity and ethnographic depth in evaluating Tanka language scholarship, particularly in a field where informal documentation remains a vital but underregulated source of evidence.

2.5. Data Extraction, Coding, and Analysis

Each selected study was coded using a structured protocol developed and refined through pilot testing to ensure consistency across heterogeneous sources. Extracted meta-data included author, year, title, document type, source, topical focus, geographic coverage, methodological approach, and publication language. All variables were organized in a spreadsheet for descriptive statistical analysis of research domains, regions, and temporal trends. A subsequent qualitative thematic synthesis identified dominant theoretical orientations, methodological tendencies, and research gaps. This mixed-methods integration provided a concise yet comprehensive overview of six decades of Tanka scholarship (see Figure 2).

No.	Author	Year	Title	Type	Source	Typical focus	Geographic focus	Method	Writing
1	Bai, Y.	2002	The phonological structure of Xiangzhou Dajia dialect	Journal Paper	Journal of Guangxi Normal University (Philosophy and Social Sciences Edition)	Phonetics and Phonology	Guangxi	Fieldwork, description	Chinese
2	Bai, Y.	2005	The phonetics of Moxian dialect of Daxu Town, of Lingnan	Journal Paper	Journal of Guilin Normal College	Phonetics and Phonology	Guangxi	Fieldwork, description	Chinese
3	Bai, Y.	2006	Oncosynonym of Dajia dialect in Xiangzhou County in Guangxi Province	Journal Paper	Journal of Guangxi Normal University (Philosophy and Social Science Edition)	Phonetics and Lexicon	Guangxi	Fieldwork, description	Chinese
4	Bai, Y.	2006	Site—The trace of the historical phonetic evolution of Xiangzhou Dajia dialect	Journal Paper	Journal of Guangxi University for Nationalities (Philosophy and Social Science Edition)	Phonetics and Phonology	Guangxi	Fieldwork, description	Chinese
5	Bai, Y.	2007	A study on the phonetics of the Guangxi Dajia dialect	Monograph	Guangxi People's Publishing House	Phonetics and Phonology	Guangxi	Fieldwork, description, comparison	Chinese
6	Bai, Y.	2007	The affiliation of the Dajia dialect in southern Guangxi	Journal Paper	Dialect	Affiliation	Guangxi	Description, comparison	Chinese
7	Bai, Y.	2007	The phonetics of Dajia dialect of Yongfeng County	Journal Paper	Journal of Guilin Normal College	Phonetics and Phonology	Guangxi	Fieldwork, description	Chinese
8	Chan, W. F.	2012	A brief introduction to the phonetics of Tanka dialect in Tai O, Hong Kong	Journal Paper	Journal of Southern Linguistics	Phonetics and Phonology	Hong Kong	Fieldwork, description, comparison	Chinese
9	Chang, S. H., & Zhuang, C. S.	2003	The dialects of the New Territories	Monograph	The Commercial Press (Hong Kong)	Phonetics and Phonology	Hong Kong	Fieldwork, description	Chinese
10	Chen, F. T.	2017	A study on the Dajia dialect in Zhaoqing, Hezhou	Thesis	Guangxi Normal University	Phonetics, Lexicon, and Grammar	Guangxi	Fieldwork, description, comparison	Chinese
11	Cheng, M. M.	2020	A study on the phonetics and grammar of the Dajia dialect in Wuzhou	Thesis	Guangxi Normal University	Phonetics and Grammar	Guangxi	Fieldwork, description, comparison, quantitative analysis	Chinese
12	Cheng, M. M., & Liu, L. Y.	2024	A study on family language use among the Tanka people in Wuzhou	Journal Paper	Chinese National Expo	Language Attitudes and Language Use	Guangxi	Questionnaire and interview	Chinese
13	Cheng, S. T.	2020	Language identity and dialect endangerment—Taking the Tanka in Hong Kong's Cheung Chau as an example	Thesis	The Education University of Hong Kong	Language Identity	Hong Kong	Fieldwork, description, comparison	Chinese
14	Fung, K. K.	2012	The phonetic system and characteristics of Tanka dialect in Shek Pai Wan of Aberdeen of Hong Kong	Journal Paper	Cantonese Studies	Phonetics and Affiliation	Hong Kong	Fieldwork, description	Chinese
15	Fung, K. K.	2012	Phonetic characteristics of the Tanka dialect in Jiahu Village, Dahanu, Huangpu, Guangdong	Journal Paper	Journal of Southern Linguistics	Phonetics and Phonology	Guangdong	Fieldwork, description	Chinese
16	Fung, K. K.	2015	The language heritage and cultural transformation of boat people communities in the Pearl River Delta	Monograph	Ten Thousand Scrolls Pavilion	Phonetics and Phonology	Pearl River Delta	Fieldwork, description	Chinese
17	Fung, K. K.	2023	Phonological study of the Cantonese dialect in Hong Kong fishing villages	Monograph	Ten Thousand Scrolls Pavilion	Phonetics and Phonology	Hong Kong	Fieldwork, description	Chinese
18	Guo, S. H.	2002	A survey report on the dialect of boat people in Macau	Thesis	Jinan University	Phonetics and Lexicon	Macau	Fieldwork, description, comparison	Chinese
19	Huang, G. F.	2024	An investigation into the Tanka language of returned overseas Chinese in Pulau Rhy Island, Qiongzhou Town, Bohai, Guangxi	Monograph	World Publishing Corporation	Phonetics, Lexicon, and Grammar	Guangxi	Fieldwork, description	Chinese
20	Huang, G. F., & Xu, D. D.	2018	The nature of the Dajia dialect of the returned overseas Chinese on Pownan Island in Qiongzhou Town, Bohai City	Journal Paper	Journal of Hezhou University	Phonetics and Phonology	Guangxi	Fieldwork, description	Chinese
21	Huang, G. G.	1990	The phonological system and characteristics of the Tanka dialect in Sanya City, Hainan Province	Conference Paper	Proceedings of the Second International Conference on Cantonese Dialects	Phonetics and Phonology	Hainan	Fieldwork, description	Chinese
22	Huang, G. G.	1991	The distribution of Chinese dialects in Sanya, Hainan	Journal Paper	Dialects	Phonetics and Phonology	Hainan	Description, comparison	Chinese
23	Huang, Z. X.	2015	A study on the vocabulary of the Dajia dialect in Nanning, Guangxi	Thesis	Guangxi Normal University	Lexicon	Guangxi	Fieldwork, description, comparison, quantitative analysis	Chinese
24	Kiu, K. L.	1984	On some phonetic characteristics of the Cantonese sub-dialect spoken by the boat people from Pu Tai Island	Journal Paper	Journal of the International Phonetic Association	Phonetics and Phonology	Hong Kong	Fieldwork, description	English
25	Ko, S. C.	2021	A study of the sub water songs of the Tanka boat-dwelling family: Language features	Thesis	The Education University of Hong Kong	Phonetics and Grammar	Hong Kong	Fieldwork, description	Chinese
26	Lok, K. H.	2010	A study of the phonetics of Cantonese dialect of Tanka people in Cheung Chau	Thesis	Hong Kong Shue Yan University	Phonetics and Phonology	Hong Kong	Fieldwork, description, comparison	Chinese
27	Luo, P. S.	2015	Comparative study of the differences between the Shau Kei Wan waterway dialect and surrounding waterway dialects in Hong Kong	Thesis	Hong Kong Shue Yan University	Phonetics and Phonology	Hong Kong	Fieldwork, description, comparison	Chinese
28	Mao, C. Z.	2015	Review of research on the language of the Dan people	Journal Paper	Theoretical Observation	Literature Review	Lingnan	Description, comparison	Chinese
29	Mao, C. Z.	2019	On language attitudes of the Dan people in Hainan from the perspective of sociolinguistics	Journal Paper	Humanities & Social Sciences Journal of Hainan University	Language Attitudes and Language Use	Hainan	Questionnaire and interview	Chinese
30	McCoy, J.	1965	The dialects of Hong Kong boat people: Kaa Sai	Journal Paper	Journal of the Hong Kong Branch of the Royal Asiatic Society	Phonetics and Lexicon	Hong Kong	Fieldwork, description, comparison	English
31	Pan, J. Y.	2001	The variation of the Raping Dajia Cantonese dialect and the loss of its communicative function	Journal Paper	Journal of Shantou University (Humanities and Social Sciences Edition)	Phonetics and Phonology	Guangdong	Fieldwork, description	Chinese
32	Pan, J. Y., & Luo, L. L.	2013	The Dajia people and dialect along the coast of Haifeng	Journal Paper	Journal of Hainan Normal University	Phonetics and Lexicon	Guangdong	Fieldwork, description	Chinese
33	Pu, I.	2019	A study on the phonetics and grammar of the boat dialect in the urban area of Guilin	Thesis	Guangxi Normal University	Phonetics and Grammar	Guangxi	Fieldwork, description, comparison	Chinese
34	Wang, C., Peng, D. X. W., Dai, Y. M. & Q. C.	2025	The endangered Tanka language in Hong Kong: Phonological variations and lexical convergence with Cantonese	Journal Paper	Humanities & Social Sciences Communications	Phonetics and Lexicon	Hong Kong	Fieldwork, description, comparison, quantitative analysis	English
35	Wu, Y. X.	2007	A comprehensive analysis of the regional characteristics of the Tai O waterway dialect	Thesis	Hong Kong Shue Yan University	Phonetics, Lexicon, and Grammar	Hong Kong	Fieldwork, description, comparison	Chinese
36	Siu, W. H.	2017	A study on the phonetics of the Raping Dajia dialect	Thesis	Guangdong Polytechnic Normal University	Phonetics and Phonology	Guangdong	Fieldwork, description	Chinese
37	Siu, W. H.	2016	A survey of language use of the Raping Dan people	Journal Paper	Journal of Hainan Normal University	Language Attitudes and Language Use	Guangdong	Questionnaire	Chinese
38	Yang, B., & Cheng, M. M.	2020	Literature review of Tanka language in Lingnan area	Journal Paper	Journal of Wuzhou University	Literature Review	Lingnan	Description and comparison	Chinese
39	Yang, J. Y.	2019	A study on the phonetics and vocabulary of the Dajia dialect in Chongkoumen fishing village, Nanhai Town, Zhongshan City	Thesis	Guangxi Normal University	Phonetics and Lexicon	Guangdong	Fieldwork, description, comparison, quantitative analysis	Chinese
40	Yao, Q. Z.	2022	Sound variations of Dajia dialect in Shatian Town, Dongguan City	Journal Paper	Journal of Guangdong Polytechnic Normal University	Phonetics and Phonology	Guangdong	Fieldwork, description	Chinese
41	Zhang, S., & Huang, X.	1988	Examination of the "Dajia" water people in the Pearl River Estuary	Journal Paper	Social Science Frontline: Ethnology	Phonetics and Phonology	Pearl River Delta	Description, comparison	Chinese
42	Zhuang, C. S.	2009	Dialects of waterborne residents (Tanka) in Lingnan area	Journal Paper	Cultural Heritage	Phonetics and Affiliation	Lingnan	Fieldwork, description, comparison	Chinese

Figure 2. Targeted Tanka language studies (Bai, 2002, 2005, 2006a, 2006b, 2007a, 2007b, 2007c; Chan, 2012; Chang & Zhuang, 2003; Chen, 2017; Cheng, 2020; Cheng & Liu, 2024; Cheung, 2020; Fung, 2012a, 2012b, 2015, 2023; Guo, 2002; G. F. Huang, 2024; G. F. Huang & Xu, 2018; G. G. Huang, 1990, 1991; Z. X. Huang, 2015; Kiu, 1984; Ko, 2021; Lok, 2010; Luo, 2015; Mao, 2015, 2019; McCoy, 1965; Pan, 2001; Pan & Luo, 2013; Pu, 2019; Wang et al., 2025; Wu, 2007; Xu, 2016, 2017; J. Y. Yang, 2019; B. Yang & Cheng, 2020; Yao, 2022; Zhang & Huang, 1988; Zhuang, 2009).

2.6. Assessment of Risk of Bias

Potential selection, publication, and methodological biases were systematically examined. Multiple databases and multilingual inclusion minimized selection and publication bias. Methodological bias was evaluated through a qualitative checklist adapted from PRISMA's risk-of-bias tool, emphasizing design transparency, representativeness, and replicability. Reviewer subjectivity was mitigated through cross-validation and consensus discussions. Overall, this ensured balanced representation and reliability in synthesizing a historically uneven research field.

3. Findings

3.1. Research Trends

3.1.1. Temporal Trends

Analysis of publications on the Tanka language over the past six decades reveals a markedly uneven temporal distribution, characterized by extended periods of scholarly inactivity interspersed with isolated studies, as demonstrated in Figure 3. Prior to the 2000s, research output was sporadic, with only incidental references appearing as single publications in 1965, the 1980s, and the early 1990s. This pattern indicates that a lack of sustained academic engagement during this early period. A modest increase in activity emerged in the early 2000s, with a notable cluster of studies between 2001 and 2006. However, consistent scholarly attention only materialized after 2007, a peak year yielding four studies. This shift coincided with growing academic interest in language contact, endangered languages, and ethnolinguistic identity in Southern China. The most significant

surge occurred post-2010, particularly during 2012–2020. This period featured repeated peaks: three studies in 2012, four studies in 2015, three studies in 2019, and three studies in 2020. This sustained growth aligns with heightened global and national focus on intangible cultural heritage preservation, alongside increased research funding and the expanded open-access platforms. Recent outputs (2023–2025) reflect moderate but steady publication rates, suggesting plateauing interest after years of growth. Collectively, these trends demonstrate Tanka language research’s revolution from marginal academic curiosity to a recognized area of scholarly inquiry, particularly in the context of sociolinguistics and language endangerment studies.

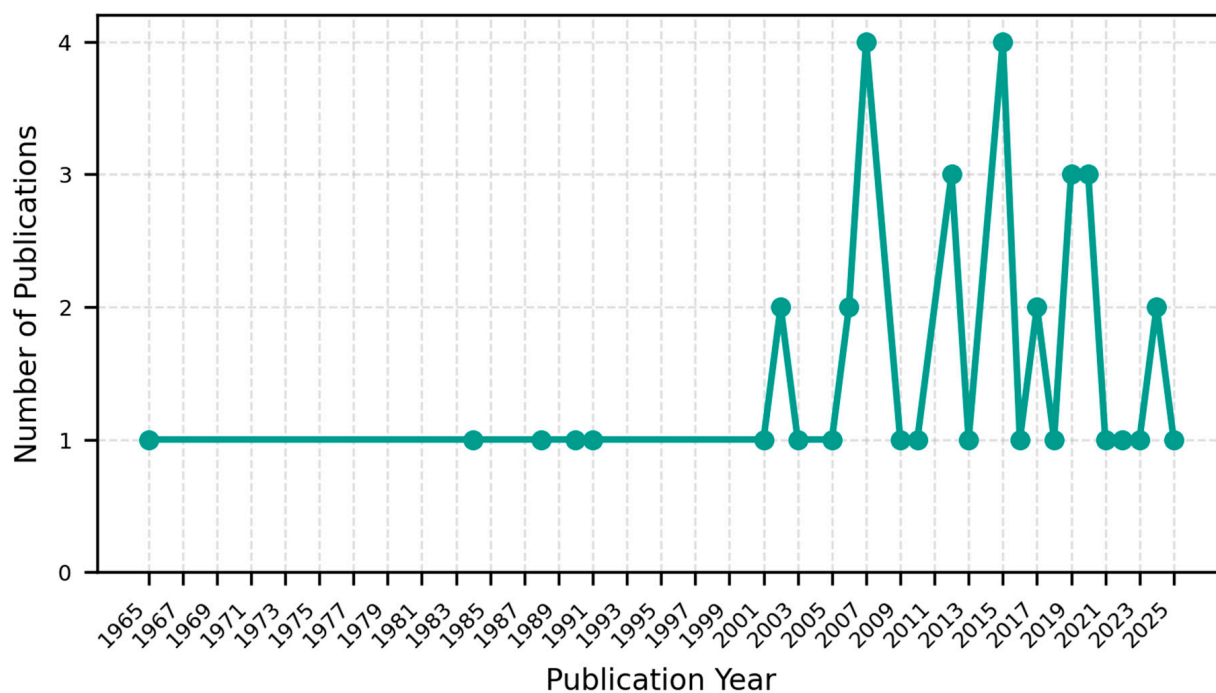


Figure 3. Number of studies published from 1965 to 2025.

3.1.2. Topical Distribution

As illustrated in Figure 4, phonetic and phonological studies dominate Tanka language research, accounting for roughly half of all publications, with another 33 % combining phonetic analysis with lexical or grammatical description and, in some cases, discussions of linguistic affiliation. This emphasis on segmental inventories, tonal patterns, and suprasegmental features, often in comparison with Cantonese, Hakka, and other Sinitic varieties, reflects both the accessibility of phonetic data in fieldwork and the long-standing dialect phonology tradition in Chinese linguistics.

In contrast, lexical and grammatical studies remain sparse, with only a handful of works addressing morphosyntax or clause structure in depth. Research on genealogical affiliation and typological classification of Tanka within the Sinitic family is similarly limited. More recently, a modest rise in studies on language identity, attitudes, and use suggests growing attention to the sociolinguistic dimensions of Tanka, though comprehensive grammatical and functional analyses remain notably underdeveloped.

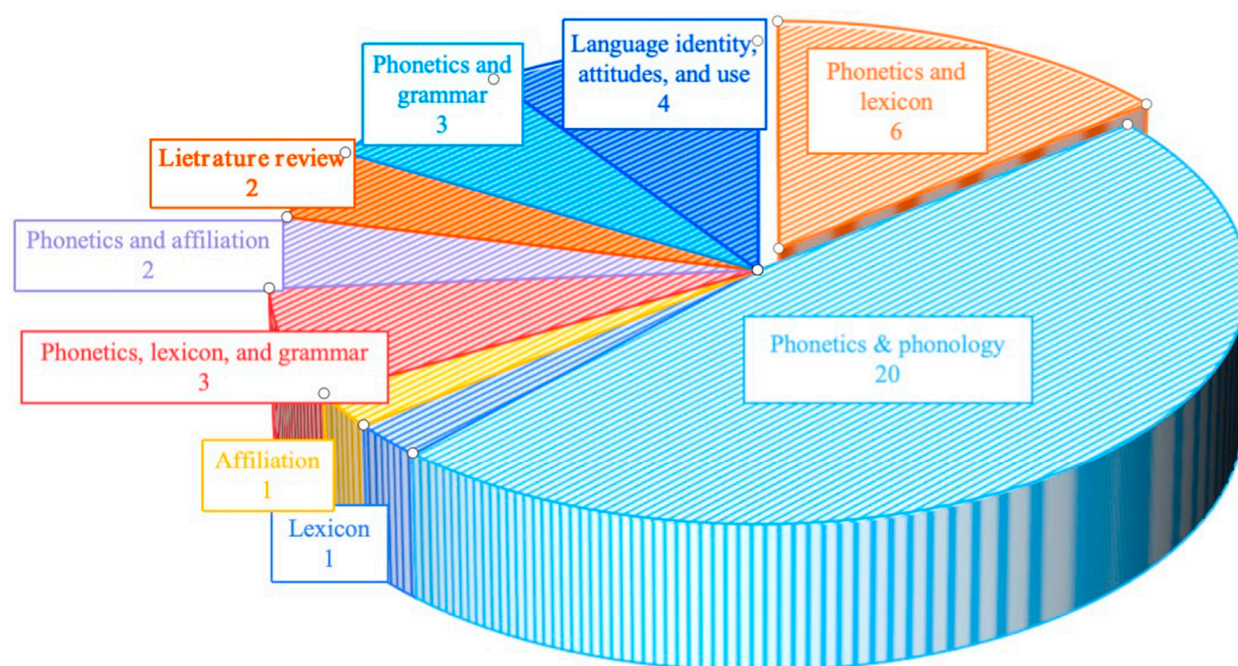


Figure 4. Distribution of research subjects.

3.1.3. Methodological Trends

Methodologically, Tanka research remains dominated by qualitative approaches, centered on descriptive linguistics and ethnographic fieldwork. As shown in Figure 5, nearly 45% (19 of 42) of studies combine fieldwork with linguistic description, focusing on phonetic transcription, lexical elicitation, and structural analysis based on primary data. About 26% (11 studies) extend this approach through comparative analysis, situating Tanka within broader Sinitic contexts. A smaller subset (12%, five studies) adopts descriptive-comparative methods using secondary data only, lacking field validation. Quantitative and computational methods remain marginal, accounting for less than 10% (four studies) that integrate statistical modeling, dialectometry, or corpus-based analysis. Similarly, sociolinguistic survey research is limited, with only 7% (three studies) employing questionnaires or interviews. Overall, methodological diversification is minimal, underscoring the need for data-driven, interdisciplinary frameworks to complement the prevailing qualitative tradition.

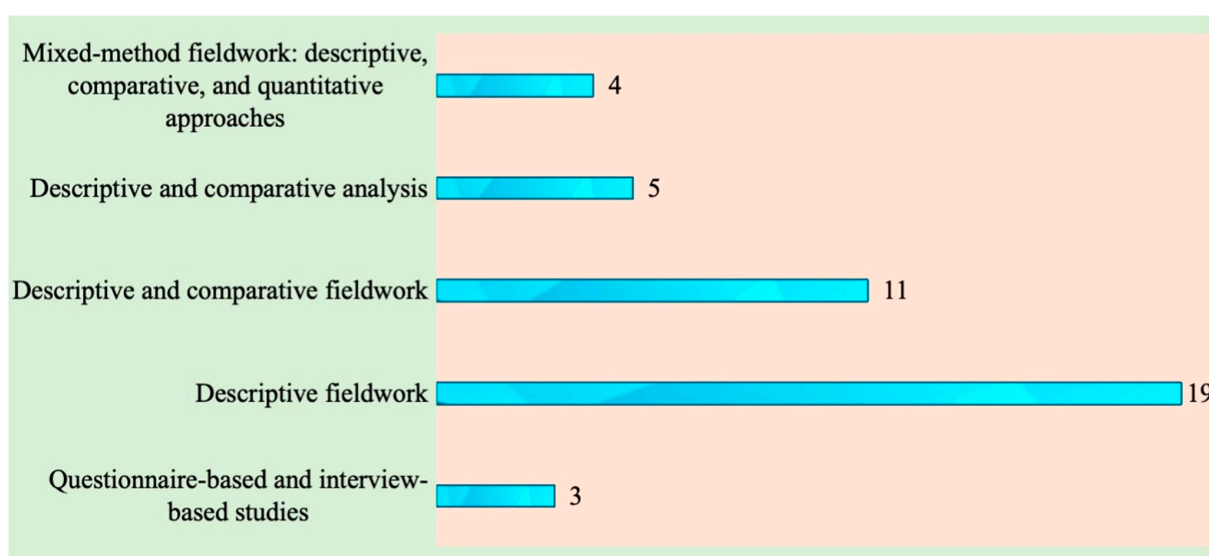


Figure 5. Distribution of research topic focus.

3.1.4. Geographic Focus

As shown in Figure 6, existing research displays a marked geographic concentration, with Guangxi (14 studies) and Hong Kong (12 studies) emerging as the primary sites for Tanka linguistic and sociocultural investigation. This reflects the historical prominence and demographic density of Tanka-speaking communities in these coastal and riverine regions. In contrast, Guangdong (seven studies), despite its centrality in the Pearl River Delta, has received comparatively less focused attention, partly because many works subsume it under broader regional labels such as “Pearl River Delta” or “Lingnan”, which obscures local variation.

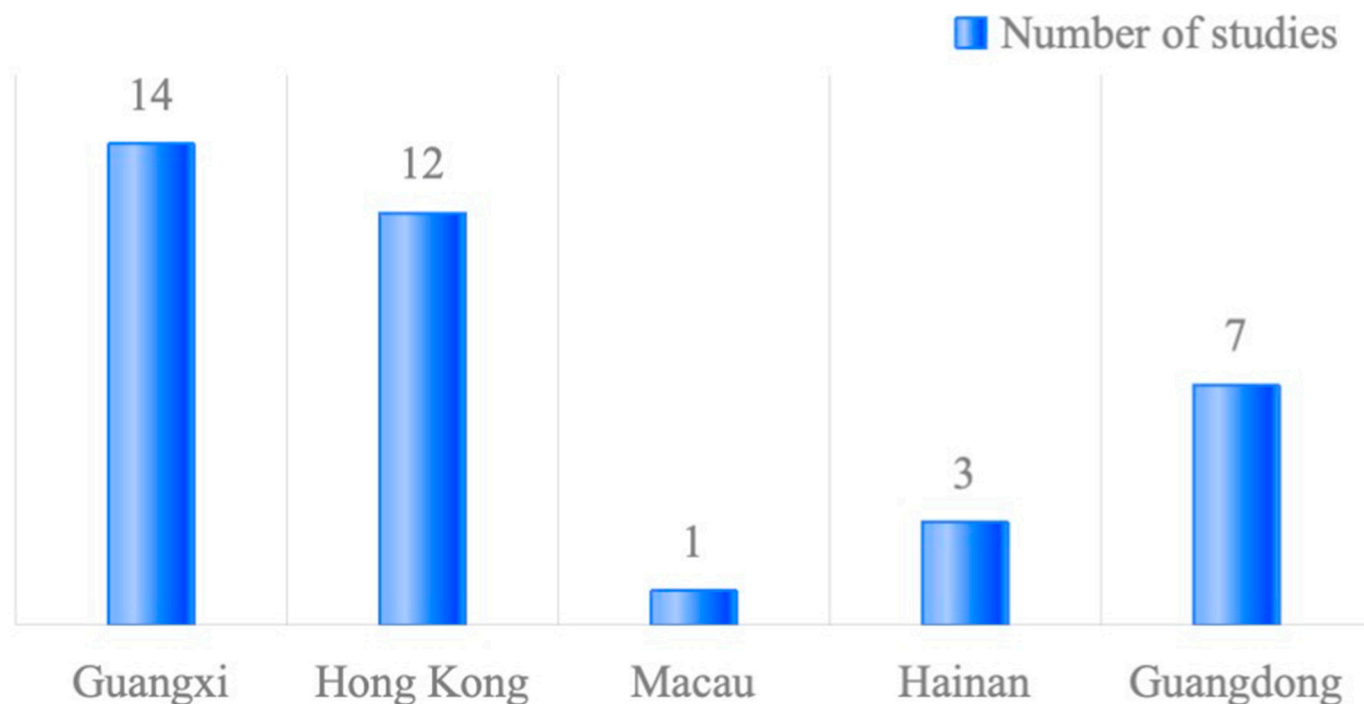


Figure 6. Distribution of Tanka language studies by geographic focus.

Significant gaps remain in Hainan (three studies) and Macau (one study), where fieldwork remains limited despite historical evidence of Tanka presence. These disparities likely stem from field access constraints, institutional priorities, and the peripheral academic visibility of these regions. The overall imbalance highlights a broader trend in which historically prominent and accessible areas (e.g., Hong Kong and Guangxi) attract disproportionate attention, while peripheral or logistically challenging regions remain underrepresented. To ensure analytical clarity and avoid data overlap, studies using broad geographic terms were excluded from statistical visualization, allowing a more precise representation of the actual research coverage.

3.1.5. Language of Publication

The distribution of publication languages reveals a pronounced imbalance. Of the 42 studies analyzed, 92.86% (39) are published in Chinese, with only 7.14% (3) in English (see Figure 7). This dominance reflects the local academic context in which Tanka research has developed—primarily within Chinese universities, regional journals, and domestic conferences. While linguistically natural, this pattern severely limits international visibility and scholarly exchange. The minimal presence of Tanka studies in English-language journals renders their findings largely inaccessible to global linguistics and anthropology communities, and the absence of Tanka data in international repositories such as ASJP

and PHOIBLE further constrains comparative research. Overcoming this barrier requires promoting bilingual publication, translation of key works, and integration of Chinese-language studies into global linguistic platforms, thereby enhancing visibility, theoretical engagement, and cross-disciplinary collaboration in Tanka research.

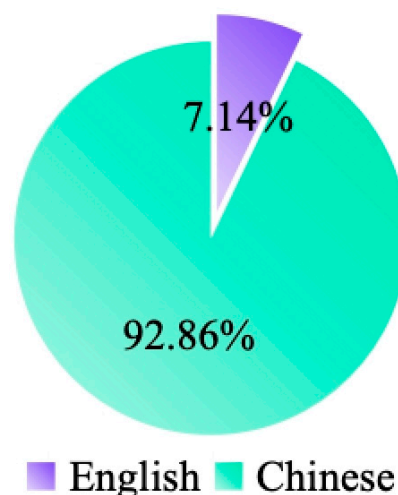


Figure 7. Distribution of language of publication.

3.2. Linguistic Structure of the Tanka Language

3.2.1. Phonetics and Phonology

The phonetic and phonological system of the Tanka language exhibits both convergence with and divergence from Standard Cantonese, reflecting prolonged bilingual contact, ecological adaptation, and sociolinguistic assimilation (Bai, 2007b; Chan, 2012; Cheng, 2020; Fung, 2012a, 2015, 2023; Guo, 2002; G. F. Huang, 2024; G. F. Huang & Xu, 2018; Wang et al., 2025). Comparative studies reveal internal variation across regions, shaped by localized ecologies and contact histories. Notably, Bai (2007a) and Pu (2019) identified strong correspondences in initials, finals, and tonal categories between Northern Guangxi Tanka and Pinghua, suggesting a shared phonological base.

A hallmark feature across many Tanka varieties is the merger of initial consonants, particularly the replacement of the alveolar nasal /n/ with the lateral /l/, as in /nam/ ('south') → /lam/. While this feature also appears in colloquial Hong Kong Cantonese, it is realized with greater consistency in Tanka speech (Bai, 2007a; Chan, 2012; Guo, 2002). Another frequently observed phenomenon is the instability of initial /ŋ/, which is often deleted in casual speech, as in 愛 /ŋoi/ ('to love') → /oi/ (Fung, 2015). These patterns reflect a combination of articulatory simplification and phonological convergence with dominant Cantonese varieties. In terms of complex onsets, labialized velar stops /kw/ and /k'w/ are typically simplified to plain velars, as in 過 guò /k'wɔ/ → /kɔ/ and 龜 guī /k'wɛi/ → /kei/. Likewise, the approximant /w/ is frequently realized as the voiceless fricative /f/, as in 湖 hú /wu:/ → /fu/ and 戶 hù /wu:/ → /fu/ (Guo, 2002; G. F. Huang & Xu, 2018). This glide-to-fricative shift likely results from historical phonologization reinforced by substrate influence or regional contact. Regional variation further complicates the phonological profile of Tanka. In areas such as Chaozhou, Shanwei, and Huidong, Tanka speakers preserve voiced initials /b/, /g/, and /z/, features characteristic of Chaoshan Minnan dialects but absent in the Tanka varieties spoken in Guangzhou and the Pearl River Delta. This suggests significant phonological borrowing driven by geographic proximity and sustained intergroup interaction (Pan, 2001; Xu, 2017).

The vowel system of Tanka shows a consistent trend of neutralizing roundedness contrasts found in Cantonese. High front rounded vowels /y/ typically shift to unrounded /i/,

as in 魚 yú /jy:/ → /ji/, while diphthongs like /œy/ and /øy/ become /ei/ or /ui/, as in 佢 /k'œy/ → /k'ei/ and 去 qù /hœy/ → /hei/ (Chan, 2012; Fung, 2012a). Mid-rounded monophthongs /œ:/ shift to /ɛ/ (e.g., 靴 xuē /hœ:/ → /hɛ/), and nasalized vowels like /œŋ/ are realized as /oŋ/ (e.g., 良 liáng /lœŋ/ → /loŋ/). These patterns suggest a broader regional tendency toward unrounding, likely influenced by articulatory ease and dialect contact. Coda simplification is another defining feature. Bilabial codas like /p/ are often dropped entirely (e.g., 鴨 yā /ap/ → /a/), while alveolar and velar codas such as /t/ and /k/ frequently undergo glottalization or deletion, as in 屋 wū /uk/ → /uʔ/ and 食 shí /sik/ → /siʔ/ (Chan, 2012; Fung, 2015). Nasal codas also undergo neutralization, with /ŋ/ often realized as /n/ (e.g., 岸 àn /oŋ/ → /on/) and /m/ as /n/ (e.g., 點 diǎn /tim/ → /tin/), reflecting contact-induced phonotactic restructuring and articulatory simplification (J. Y. Yang, 2019; Wang et al., 2025).

Tonal systems vary widely (five to nine tones) (Bai, 2007a, 2007b; Chan, 2012; Fung, 2015). Some communities retain entering tones (e.g., 搭 dā /taʔ/), while others show tone mergers and splits conditioned by voicing or vowel quality. Suprasegmentally, Tanka is marked by slower tempo, creaky voice, and distinctive intonation contours, particularly in ritual and narrative contexts (Cheng, 2020; Fung, 2012a, 2023).

Overall, the Tanka phonological system reflects a hybrid equilibrium, maintaining conservative features alongside innovations arising from ecological adaptation and long-term contact. These phonetic shifts, beyond structural change, function as sociophonetic markers of identity, reinforcing Tanka's distinctiveness within the southern Sinitic continuum.

3.2.2. Lexicon

The Tanka lexicon embodies the community's aquatic livelihood, encoding maritime knowledge, ecological adaptation, and cultural identity distinct from agrarian-based Sinitic lexicons (Chen, 2017; Fung, 2015; Guo, 2002; G. F. Huang, 2024; G. F. Huang & Xu, 2018; Wang et al., 2025; J. Y. Yang, 2019). Fishing terminology differentiates sharply between marine and riverine ecologies: expressions such as 做海 zuòhǎi /tsou³² hoi³⁵/ 'to make a living by fishing at sea', 打山口 dǎ shānkǒu /ta³² san⁴⁴ hœu³⁵/ 'to set a net to fish at a recognizable mountain spot without buoys to avoid theft', and 落罟 luògū /loʔ³² ku⁴⁴/ 'to set fixed nets for fishing' reflect detailed knowledge of tidal rhythms and fish migration patterns (Fung, 2015; Guo, 2002; G. F. Huang & Xu, 2018; Wang et al., 2025), while riverine terms such as 拉纜 lālǎn 'to pull mooring lines for docking or maneuvering a boat' and 拋江 pāojiāng 'the practice of rowing the boat laterally across a river' illustrate adaptation to freshwater contexts (Z. X. Huang, 2015).

Tanka's folk taxonomy demonstrates remarkable ecological precision, distinguishing aquatic species by habitat, morphology, and seasonality, often exceeding scientific classifications (Chen, 2017; Z. X. Huang, 2015). Similarly, boat-related vocabulary encodes both technological and social practices. Terms such as 火艇 huǒtǐng /fo³⁵ t'ien²⁴/ 'boat used for illumination when fishing', 拖蝦船 tuōxiā chuán /t'ɔ⁴⁴ ha⁴⁴ sin²¹/ 'shrimp trawler', 歌堂船 gētáng chuán /kɔ⁴⁴ t'ɔŋ²¹ sin²¹/ 'boat for hosting banquet', 雪艇 xuětǐng /siʔ³² t'ien²⁴/ 'Ice transport vessel' exemplify not only practical seafaring knowledge but also the cultural transmission of specialized marine craftsmanship through generations (Chen, 2017; Fung, 2015; Guo, 2002; Z. X. Huang, 2015; Wang et al., 2025). Despite regional variation, a core set of terms, e.g., 住家艇 zhùjiā tǐng /tsi³² ka⁴⁴ t'ien²⁴/ ('houseboat') and 生艙 shēngcāng /san⁴⁴ ts'ɔŋ⁴⁴/ 'a live fish holds or live well compartment on fishing vessels', is shared across marine and riverine communities, underscoring occupational continuity and cultural cohesion (Chen, 2017; Fung, 2015; Guo, 2002; Z. X. Huang, 2015; Wang et al., 2025; J. Y. Yang, 2019).

Morphologically, Tanka shares compounding, reduplication, and affixation with Cantonese but diverges through its rich use of bound morphemes such as the diminutives (e.g., -仔 *zǎi*), and aspect markers (e.g., -咗 *zuo*), often embedded in maritime contexts (Guo, 2002; Z. X. Huang, 2015; Wang et al., 2025). This reflects both convergence with Cantonese and typological conservatism distinct from Mandarin, situating Tanka within the southern Sinitic sphere.

Etymologically, the lexicon reveals three historical strata: (1) Archaic Sinitic forms, including ritual and taboo vocabulary (Z. X. Huang, 2015); (2) Borrowings from Cantonese and Hakka, especially in basic and functional lexicon (Guo, 2002); and (3) Zhuang-Dong substrata in kinship and animistic terms, particularly in coastal Guangxi and Hainan (Chen, 2017; Wang et al., 2025). This multilayered structure underscores Tanka's adaptive resilience and its role as a dynamic linguistic interface shaped by sustained maritime contact.

3.2.3. Grammar

Compared with phonological and lexical studies, grammatical research on the Tanka language remains limited (Chen, 2017; Cheng, 2020; G. F. Huang, 2024; Pu, 2019). Existing analyses nonetheless reveal significant syntactic convergence with neighboring Sinitic varieties, particularly Pinghua and Cantonese. Pu (2019) provides a comprehensive account of the Guilin Tanka variety, identifying strong structural parallels with Pinghua. Both employ yes–no interrogatives in alternative question formation and share 把 *bǎ*-like disposal constructions that differ only in preposition choice. The passive marker /ta³⁵/, historically meaning 'to give', has grammaticalized into a causative and subsequently passive marker, a pathway typical of southern Sinitic languages. In comparatives, Guilin Tanka favors 比 *bǐ* over 过 *guò*, unlike Cantonese. Flexible adverbial placement (先 *xiān* 'first', 多 *duō* 'much', and 晚 *wǎn* 'late' following verbs) further aligns Guilin Tanka with Pinghua typology while marking local innovation. Cheng (2020), analyzing Wuzhou Tanka, identifies similar areal alignment with Guangfu Cantonese. Core constructions include disposal constructions (marked by 拧 *nǐng*), passives (introduced by 挨 *ái*, 着 *zhuó*, or 畀 *bì*), and double-object and existential sentences. Interrogatives appear as yes–no, alternative, A-not-A, and wh-forms, typically ending with 咧 *liē* or 嘅 *ge*, while 啊 *a* and 啦 *la* are rare. Imperatives center on action predicates with optional subjects. Together, these findings indicate that Tanka grammar reflects deep areal diffusion, integrating Pinghua-like syntactic structures with Cantonese typological features. Yet the overall grammatical system remains underdocumented, calling for expanded comparative and functional analyses to clarify its typological status within the Sinitic family.

3.3. Genetic Affiliation

The genetic affiliation of Tanka varieties reveals a heterogeneous linguistic landscape shaped by historical contact, migration, and sociopolitical marginalization (Bai, 2007b; Chan, 2012; G. F. Huang & Xu, 2018; McCoy, 1965; Pu, 2019; Wang et al., 2025). In the Pearl River Delta (e.g., Hong Kong, Zhuhai, and Jiangmen), Tanka speech aligns closely with Cantonese, sharing core phonological, lexical, and syntactic features (Chang & Zhuang, 2003; Xu, 2017; Yao, 2022). Yet distinct micro-variations persist, such as archaic vocabulary, phonetic simplification, and occupationally specialized lexicon, reflecting Tanka's semi-nomadic maritime heritage (Pan, 2001; Z. X. Huang, 2015; B. Yang & Cheng, 2020). Some scholars highlight syntactic distinctions in sentence-final particles (Ko, 2021) and phonological divergence from Cantonese, particularly in rhyme and tonal evolution (Wang et al., 2025), suggesting partial retention of pre-Cantonese features. In northern Guangxi, especially around Guilin, Tanka varieties exhibit stronger affinities with Pinghua, shar-

ing tonal systems, lexicon, and morphosyntax shaped by long-term multilingual contact among Pinghua, Cantonese, and Hakka (Bai, 2007c; Pu, 2019).

Overall, Tanka varieties do not form a single genealogical branch but a contact-induced convergence zone influenced by Cantonese and Pinghua. Such convergence arises through prolonged bilingualism and sociolinguistic accommodation (Thomason, 2001), producing localized varieties whose structures reflect both ecological adaptation and social interaction. Understanding Tanka's position within the Sinitic family thus requires a contact-oriented, areal perspective rather than a purely genealogical one.

3.4. Language Identity, Attitudes, and Use

Recent studies on Tanka-speaking communities in Wuzhou (Cheng & Liu, 2024), Hainan (Mao, 2019), and Raoping (Xu, 2016) consistently reveal a pattern of language shift, identity re-negotiation, and intergenerational transmission loss within historically marginalized boat-dwelling groups. Across all regions, Tanka use has sharply declined across generations. In Wuzhou, Cheng and Liu (2024) document a five-generation family's transition from Tanka-dominant communication to Cantonese–Mandarin code-switching, culminating in the near extinction of Tanka among younger speakers. Similarly, in Hainan, Mao (2019) reports a decline from 97% to 43% in family-domain use, with Mandarin increasingly replacing Tanka in both domestic and public contexts due to education and state policy. The Raoping case represents the extreme: Tanka has almost disappeared, supplanted by Chaoshanese through geographic relocation, intermarriage, and occupational integration (Xu, 2016). Collectively, these findings confirm the Tanka language's critically endangered status across southern China.

Attitudinal studies reveal deep ambivalence. Elders value Tanka as a symbol of heritage yet discourage its transmission to ensure socio-economic mobility via Mandarin (Cheng & Liu, 2024). In Hainan, Mao (2019) finds that positive affective attachment rarely translates into actual use; pragmatic priorities dominate. Similarly, Xu (2016) notes that Raoping speakers increasingly perceive Tanka as socially obsolete, favoring more prestigious regional languages. These findings reflect a broader identity tension between cultural loyalty and modernization pressures. Four interconnected forces underlie Tanka language erosion: (1) intermarriage, introducing dominant languages into family domains; (2) occupational transition and urbanization, shifting communication toward land-based communities (Xu, 2016); (3) Mandarin-focused education policies, narrowing functional use (Mao, 2019); and (4) enduring social stigma, whereby the “boat people” identity remains linked to low prestige and exclusion (Cheng & Liu, 2024; Xu, 2016). These studies portray Tanka language decline as a multidimensional process rooted in sociopolitical marginalization, economic transformation, and internalized linguistic hierarchies. While older generations retain symbolic attachment, pragmatic and ideological pressures have driven a collective move toward dominant languages. The Tanka case thus exemplifies how language endangerment intertwines with identity, mobility, and power, situating it within broader discussions of cultural resilience and linguistic justice in southern China.

4. Discussion

4.1. Reflections on the Development of Tanka Studies

The evolution of Tanka research over six decades mirrors both changing academic priorities and the broader interplay between linguistic scholarship and state ideology. The sporadic attention before 2000 and its revival after 2010 align with the global endangered-language movement and China's Intangible Cultural Heritage policies. Yet such recognition often folklorizes linguistic diversity, valorizing cultural display while neglecting intergenerational transmission, illustrating what May (2012) terms linguistic tokenism within

a system of cultural governance (Yeh, 2007). The field's focus on phonetics and phonology reflects a persistent structural bias in Sinitic linguistics, marginalizing morphosyntax, pragmatics, and sociolinguistics. Guided by methodological nationalism (Wimmer & Glick Schiller, 2002), many studies privilege descriptive alignment with state narratives of unity, overlooking issues of power and marginalization that define Tanka's hybrid evolution through maritime contact.

Methodologically, Tanka research remains largely qualitative and descriptive, seldom incorporating quantitative sociolinguistics, corpus methods, or phylogenetic modeling. The lack of vitality metrics (UNESCO, 2003) and ecological approaches (Haugen, 1972; Mufwene, 2001) restricts its analytical reach. Similarly, geographic coverage remains uneven—Hong Kong and Guangxi dominate, while Hainan, Guangdong, and Macau remain understudied, leaving the transregional continuum across Southeast China largely unexplored. Linguistic research on Tanka also suffers from epistemic insularity: over 90% of studies are published in Chinese, limiting global visibility and inclusion in repositories such as PHOIBLE or ASJP. This linguistic barrier reflects deeper inequities in global knowledge production. To advance the field, scholars must adopt interdisciplinary, critical paradigms that engage with symbolic violence (Bourdieu, 1991) and linguistic citizenship (Stroud, 2001). Revitalization should resist folklorism and instead empower speakers through education, community agency, and equitable representation of their linguistic heritage.

4.2. Linguistic Convergence, Contact, and Typological Boundaries in Tanka

4.2.1. Phonological Shifts and Sociophonetic Adaptation

Tanka phonology reveals a dynamic system shaped by ecological pressures, contact, and identity performance. Rather than uniform erosion, it reflects functional adaptation and symbolic differentiation, aligning with contact and ecological linguistics (Thomason & Kaufman, 1988; Haugen, 1972; Mühlhäusler, 1996). Ecological explanations, such as the proposal that noisy marine environments may favor open vowels or reduced codas to enhance audibility, have been advanced in the literature but remain heuristic rather than demonstrably causal, given the lack of targeted phonetic and perceptual studies on Tanka speech. Accordingly, such explanations should be treated as suggestive frameworks rather than established mechanisms. Processes such as the simplification of rounded vowels (/y/, /ø/) and the weakening or loss of final stops are better situated within broader, cross-linguistically attested tendencies whereby marked segments are particularly vulnerable under conditions of transmission stress and sustained contact (Mühlhäusler, 1996). From this comparative perspective, the relatively frequent unrounding and coda reduction observed in Tanka need not be attributed exclusively to ecological pressures but can be understood as part of wider Sinitic and areal patterns of phonological change.

Intergenerational standardization in Hong Kong and Macau demonstrates convergence toward Cantonese, reflecting Fishman's (1991) domain theory: heritage phonology retreats to familial or ritual contexts. Education and media promote phonological leveling, while younger speakers symbolically retain traditional tones for nostalgic or ritual identity (UNESCO, 2003; Grenoble & Whaley, 2006). Regional variation further evidences contact-induced diffusion: voiced stops in Raoping and tonal patterns in Guilin reveal convergence with Min and Pinghua (Pu, 2019; Xu, 2017). These align with Aikhenvald's (2007) and Enfield's (2005) models of linguistic area interaction, whereby local bilingualism drives phonological hybridity rather than uniform typological change.

Finally, phonological variation in Tanka may carry indexical meanings, though the extent and stability of such meanings remain to be empirically established. The merger of /n/ and /l/, observed in several Tanka varieties, is a well-attested change across many

Sinitic languages and dialects, including varieties of Cantonese in Hong Kong, and therefore cannot be treated as a Tanka-specific innovation. While mergers of this type have been interpreted elsewhere in terms of sociophonetic stratification and contact-induced change (Trudgill, 1986), the currently available evidence for Tanka is largely descriptive. Determining whether comparable sociophonetic mechanisms operate in Tanka will require systematic, comparative sociolinguistic research that examines parallel developments in Cantonese and other regional Sinitic varieties (Fung, 2015). Nevertheless, existing observations suggest that Cantonese-like phonological realizations are often associated with education, mobility, and social integration, whereas features such as creaky voice quality or the retention of entering-tone reflexes tend to be associated with heritage-oriented or intra-community contexts. Interpreted cautiously, this pattern is consistent with indexical approaches to variation, in which phonological features simultaneously serve communicative functions and index social meanings (Eckert, 2008; Silverstein, 2003).

4.2.2. Contact-Ecological Dynamics of the Tanka Lexicon

The Tanka lexicon constitutes a repository of maritime epistemology, reflecting ecological adaptation and layered contact. Core semantic domains, boats, fishing, aquatic species, encode experiential knowledge and cultural semiotics. Terms such as 打山口 *dǎ shānkǒu* /tā³² san⁴⁴ hēu³⁵/ ‘setting nets without buoys at landmark spots’ or 火艇 *huǒtǐng* /fǒ³⁵ t’ien²⁴/ ‘fishing illumination boat’ exemplify the co-grammaring of environment and practice (Evans, 2010).

Morphologically, the Tanka language shares typological features with Cantonese, including compounding, reduplication, and affixation, yet it deploys them within distinctive sociocultural frames. Morphemes like -仔 *zǎi* (diminutive) and -㗎 *zo*, (aspectual marker) persist not only due to grammatical function but also through cultural embedding. Their presence aligns with Aikhenvald’s (2007) notion of peripheral morphological diffusion under contact and shows both retention and refunctionalization. The Tanka use of these morphemes often indexes social roles (e.g., kinship diminutives) or those temporality tied to maritime routines, suggesting selective specialization rather than general borrowing. Etymologically, its strata combine archaic Sinitic, regional borrowings (Cantonese, Hakka), and Zhuang–Dong substrata, forming a multilayered lexicon of contact (Enfield, 2005; Silverstein, 2003). Each layer bears ideological significance: archaic forms connote ritual prestige, while loans reflect pragmatism and adaptability. Lexical retention now serves symbolic and performative roles. Words like 住家艇 *zhùjiā tǐng* /tsi³² ka⁴⁴ t’ien²⁴/ ‘houseboat’ or 生艚 *shēngcāng* /san⁴⁴ ts’ən⁴⁴/ ‘live-fish hold’ function as semiotic emblems of Tanka identity (Gal & Irvine, 1995; Agha, 2007). Despite language shift, emblematic vocabulary sustains collective memory and sociocultural value.

From a documentation perspective, the lexicon should form the core of revitalization efforts (Himmelman, 1998). Priorities include thematic lexical archiving (boats, fishing, aquatic biodiversity), integration into education and digital platforms, and participatory documentation ensuring community agency (Austin & Sallabank, 2011). Rather than a passive remnant, vocabulary becomes an active site of resilience and identity reconstruction.

4.3. Genealogical Complexity and Contact-Induced Variation

The genetic affiliation of the Tanka language complex defies linear genealogical classification. Rather than forming a unified Sinitic sub-branch, Tanka exemplifies contact-induced change (Thomason & Kaufman, 1988; Thomason, 2001), whereby linguistic structures evolve through sustained multilingual interaction, bilingualism, and sociohistorical adaptation. Across regions, Tanka varieties display gradient affiliations. In the Pearl River Delta, Hainan, and southern Guangxi, convergence with Cantonese is evident in phonol-

ogy and syntax, while the retention of archaic lexemes and occupation-specific vocabulary reflects substrate persistence and selective innovation (Heine & Kuteva, 2005). Conversely, northern Guangxi Tanka shows alignment with Pinghua in tonal and morphosyntactic features, supporting Aikhenvald and Dixon's (2001) cline model of gradual affinity in contact zones.

To capture these dynamics, we propose a Contact–Ecological Network Model (see Figure 8), synthesizing Thomason's (2001) contact theory, Enfield's (2005) areal linguistics, and Mufwene's (2001) ecological framework of language evolution. This model conceptualizes Tanka as a convergence zone within overlapping Cantonese–Pinghua ecologies, conditioned by mobility, marginality, and sociohistorical isolation. It aligns with the wave model (Schmidt, 1872) and network diversification model (François, 2014), emphasizing diffusion and shared innovations over hierarchical descent. Linguistic features are understood to spread unevenly across communities and subsystems, conditioned by intensity of contact, communicative function, and social valuation. Phonological features, for example, diffuse relatively rapidly due to their high perceptual salience and articulatory economy, as reflected in widespread vowel unrounding and coda reduction in southern and coastal Tanka varieties (Chan, 2012; Fung, 2015). In contrast, morphosyntactic features, such as passive constructions and word-order preferences, exhibit slower and more selective diffusion, with convergence observed only in specific constructions and regions, as documented in the Guilin and Wuzhou varieties (Pu, 2019; Cheng, 2020). Lexical domains show yet another pattern: while extensive borrowing from Cantonese and Pinghua is evident in general vocabulary, maritime and occupation-specific lexicon, along with certain domain-specific morphological patterns, are selectively retained across regions, reflecting the enduring role of livelihood, cultural practice, and identity in shaping linguistic persistence (Guo, 2002; Z. X. Huang, 2015; Wang et al., 2025).

A central insight emerging from this synthesis is the role of linguistic fluidity. Linguistic fluidity refers here to structured, socially meaningful variability across linguistic domains, rather than instability, incomplete transmission, or transitional competence (Agha, 2007; Silverstein, 2003). In line with sociolinguistic approaches that treat variation as indexical and socially mediated (Agha, 2007; Silverstein, 2003), linguistic fluidity captures speakers' capacity to mobilize different linguistic resources across contexts, domains, and interactional settings under sustained contact. For example, the same speakers may employ Cantonese-like phonological realizations in public or intergroup interactions, while retaining Tanka-specific lexicon and prosodic features in familial, ritual, or occupational settings (Cheng, 2020; Fung, 2015). Similarly, the coexistence of Cantonese-derived aspect markers (-咗 *zuo*) with locally specialized morphological usage in maritime discourse illustrates domain-specific accommodation rather than wholesale grammatical replacement (Guo, 2002; Z. X. Huang, 2015; Wang et al., 2025). These patterns support a view of Tanka multilingualism as networked and fluid, rather than transitional or deficient. Linguistic fluidity enables speakers to navigate multiple normative systems, producing shared innovations across contact zones while maintaining socially salient distinctions. In this sense, Tanka does not represent a failed or incomplete shift to Cantonese or Pinghua, but a stable outcome of long-term contact shaped by asymmetric power relations, consistent with network-based models of language change and diffusion (Blommaert, 2010; François, 2014).

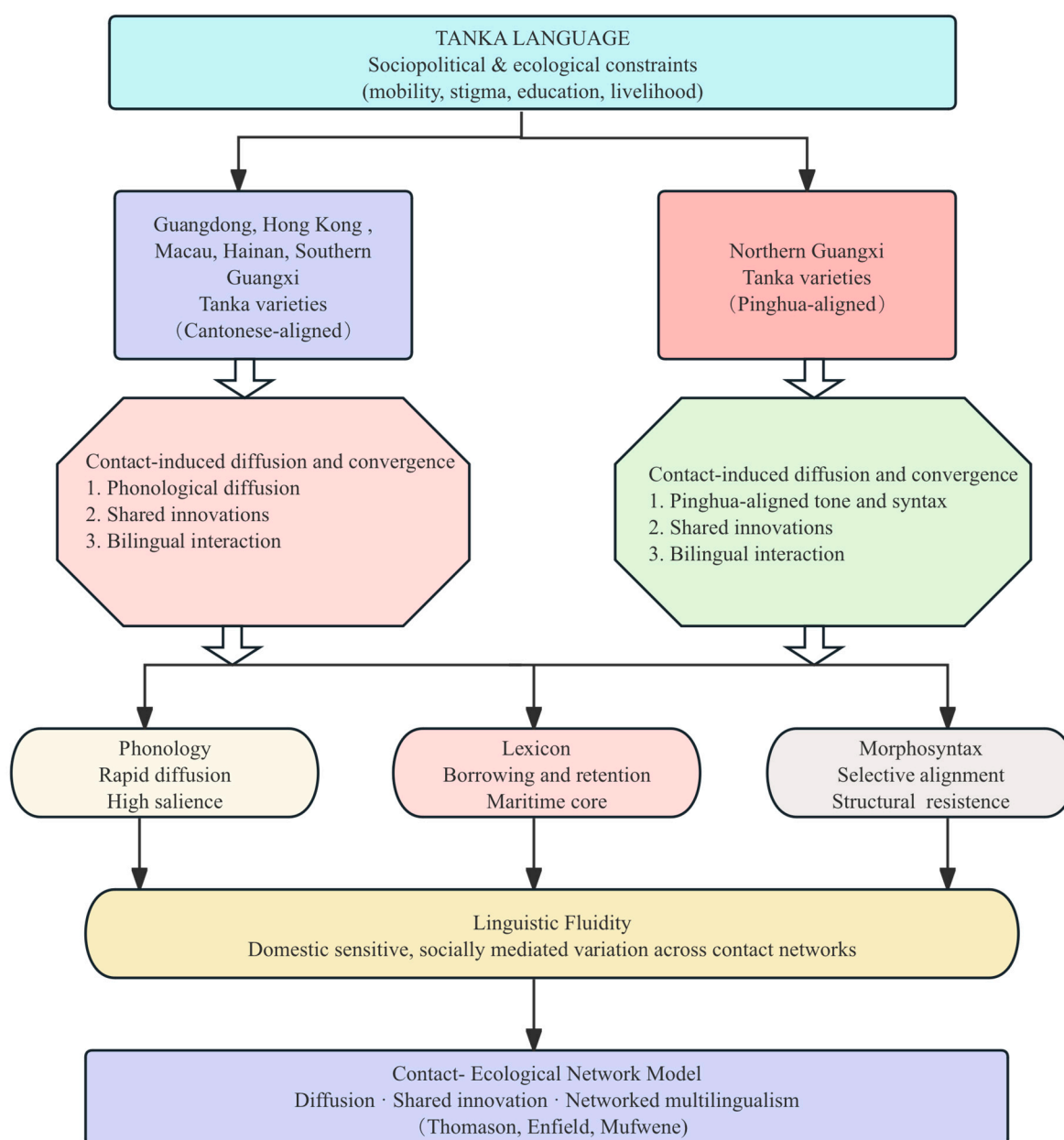


Figure 8. Tanka language contact model.

Finally, to situate this analysis within a broader theoretical frame, we propose a heuristic triangular conceptualization of multilinguality, defined by the interaction of structural constraints, linguistic fluidity, and social justice. Structural constraints delimit the space of possible variation within Sinitic grammar; linguistic fluidity captures contact-driven adaptability across domains; and social justice foregrounds the sociopolitical conditions, stigma, exclusion, and limited institutional support that shape language valuation and transmission. For example, the retreat of Tanka phonological features to elder or ritual speech (Cheng & Liu, 2024) reflects not grammatical impossibility, but socially conditioned restriction of domains. This framework does not posit Universal Grammar as a direct explanatory mechanism but treats structural universals as boundary conditions within which contact-driven and socially mediated variation unfolds.

4.4. Negotiating Language Identity, Attitudes, and Usage: Reflections on Shift and Continuity

The sociolinguistic profile of the Tanka language reveals a community navigating identity under the intertwined pressures of state standardization, linguistic hierarchies, and market rationality. Theoretically, this dynamic aligns with language ecology (Haugen, 1972; Mufwene, 2001), linguistic imperialism (Phillipson, 1992), and symbolic violence (Bourdieu, 1991), illustrating how marginalized groups internalize dominant ideologies that devalue their linguistic heritage. Across Wuzhou, Hainan, and Raoping, studies document a sharp decline in intergenerational transmission (Cheng & Liu, 2024; Mao, 2019; Xu, 2016). Functional contraction from intra-family use to symbolic retention exemplifies Fishman's (1991) disruption model, where educational Mandarinization and occupational diversification accelerate linguistic loss. Heritage domains persist only among elders or in ritualized contexts, signaling advanced language endangerment.

The existing studies uniformly reveal ambivalent attitudes among Tanka speakers. While elders view Tanka as a marker of heritage and ethnic distinctiveness, socio-economic pragmatism often supersedes cultural attachment. This mirrors broader minority language contexts where positive affect does not equate to active usage (Baker, 1992). Instead, language preference aligns with perceived utility in that Mandarin is favored as the vehicle for education, employment, and social mobility. This duality reflects Giles et al.'s (1977) Ethnolinguistic Identity Theory, where language behavior depends on perceived in-group vitality and status relations. In Wuzhou and Hainan, practical assimilationism dominates language choice (Cheng & Liu, 2024; Mao, 2019), while in Raoping (Xu, 2016), Tanka's symbolic capital is largely lost. Such dynamics also demonstrate the impact of linguistic imperialism (Phillipson, 1992), where Mandarin's dominance is naturalized through education and media, marginalizing regional and heritage languages within the state's ideological framework of national unity and modernity.

Long-standing stigma associated with "boat people" status contributes to internalized linguistic inferiority. Following Bourdieu (1991), this misrecognition of linguistic subordination perpetuates withdrawal from heritage practices. The abandonment of Tanka thus reflects not only contact-induced change but also the sedimented effects of social exclusion and symbolic violence. Language loss arises from intersecting factors, such as marital exogamy, occupational change, education policy, and urbanization, creating a self-reinforcing loop of reduced use and declining prestige. This process accords with Mufwene's (2001) ecological model: as the sociocultural environment that sustained Tanka collapses, the language loses its functional "fitness" within the communicative ecosystem. Endangerment extends beyond speaker decline to encompass identity erosion and domain displacement. According to UNESCO's (2003) vitality framework and Grenoble and Whaley (2006), effective revitalization requires restoring social functions and reversing stigma. Sustainable strategies include (a) revalorizing Tanka identity as cultural pride (Leeman, 2015); (b) integrating the language into heritage education (Hornberger, 2008); (c) promoting community-led revitalization with local agency (Leonard, 2011); and (d) leveraging digital tools such as Text-to-speech (TTS), automatic speech recognition (ASR), and large language models (LLMs), and corpus development to enhance accessibility and prestige (Bird, 2022; Arppe et al., 2016). The Tanka case exemplifies how language endangerment reflects politically mediated ecological transformation, not natural decline. Following Haugen (1972) and Mufwene (2001), language survival depends on adaptive capacity within shifting socio-environmental systems. Tanka's displacement, shaped by economic rationality and institutional ideology, underscores the need for critical, intersectional approaches that view language as both a site of marginalization and a resource for cultural resistance and empowerment.

5. Conclusions

This systematic review synthesizes six decades of research on the Tanka language, revealing both the richness of its phonological and ethnolinguistic documentation and the enduring limitations that restrict its theoretical breadth and applied impact. As a historically marginalized Sinitic variety shaped by aquatic livelihoods, mobility, and long-term contact, Tanka occupies a distinctive position in the sociolinguistic ecology across South-east China. Its linguistic profile reflects both convergence with regional prestige varieties such as Cantonese and the retention of phonological, lexical, and morphosyntactic features grounded in maritime knowledge systems. Sociolinguistic evidence consistently shows erosion of intergenerational transmission, shift toward dominant languages, and ambivalent attitudes, a pattern paralleling global endangered-language trends (Fishman, 1991; Grenoble & Whaley, 2006), yet rooted in local histories of exclusion and economic marginalization (Hornberger, 2008).

Despite valuable progress, current research remains thematically and methodologically uneven. Phonology continues to receive disproportionate attention, while morphosyntax, discourse, pragmatics, and naturalistic language use remain critically underexplored. Lexical studies often emphasize maritime cultural domains without systematically examining borrowing pathways, semantic change, or lexical diffusion. Sociolinguistic research remains fragmented and largely synchronic, lacking longitudinal depth and cross-community comparison. In terms of methodology, the field relies predominantly on qualitative field description. Limited uptake of corpus linguistics, dialectometry, vitality assessment (UNESCO, 2003), and computational analysis restricts replicability and integration into global comparative resources. The scarcity of standardized data collection protocols and open-access corpora further constrains cross-site comparability, while geographic concentration in Hong Kong and Guangxi leaves Hainan, Guangdong, Macau, and other coastal communities substantially underrepresented. Additionally, the dominance of Chinese-language scholarship restricts international visibility. Promoting bilingual or English-language publication is therefore essential not only to broaden accessibility but also to reduce epistemic asymmetries in global linguistic research and enable Tanka to participate in wider comparative, policy-oriented, and revitalization dialogues.

Future research must adopt interdisciplinary, quantitative, and community-centered frameworks. Expanding inquiry into morphosyntax, pragmatics, discourse structures, and contact typology will enhance theoretical integration within Sinitic linguistics. The development of standardized elicitation protocols, open-access corpora, and transparent statistical reporting (e.g., effect sizes, confidence intervals) will improve reproducibility and comparability (Tricco et al., 2018). One particularly promising frontier lies in AI-assisted revitalization. TTS, ASR, and LLMs can facilitate transcription, vocabulary learning, and immersive pedagogy (Meyer et al., 2022), while generative and interactive AI, including digital humans, can support culturally grounded learning environments (Bird, 2022; Ward et al., 2022). Multilingual natural language processing frameworks such as Common Voice and MASSIVE demonstrate the feasibility of community-based fine-tuning for underrepresented languages (Bapna et al., 2022), transforming documentation into active linguistic agency.

Equally critical is the development of structurally equitable, community-led revitalization models. Speakers should be positioned as co-researchers, teachers, and cultural custodians. Successful precedents such as Kōhanga Reo in New Zealand and Sámi immersion programs in Scandinavia (Austin & Sallabank, 2011; Higgins & Rewi, 2014) show how revitalization thrives when aligned with lifeways, land-based practices, and intergenerational transmission. Comparisons with maritime communities such as the Bajau Laut

further highlight the role of ecological displacement and sociopolitical hierarchies in shaping revitalization outcomes (Blust, 2006).

Ultimately, preserving Tanka is not solely a linguistic task but a commitment to epistemic justice, cultural continuity, and self-determined futures. Aligning with SDG 4 (Quality Education) and SDG 10 (Reduced Inequalities) (UNESCO, 2003; United Nations, 2015), sustainable revitalization must integrate inclusive policy support, shared research infrastructures, and community-directed pedagogical innovation. The survival of Tanka represents not only the maintenance of linguistic diversity but also the safeguarding of maritime knowledge, cultural memory, and the adaptive resilience of one of southern China's most historically enduring communities.

Author Contributions: Conceptualization, Y.D. and C.W.; methodology, Y.D. and C.W.; software, Y.D.; validation, C.W.; formal analysis, Y.D. and C.W.; investigation, Y.D. and C.W.; resources, Y.D. and C.W.; data curation, Y.D.; writing—original draft preparation, Y.D.; writing—review and editing, C.W. and Y.D.; visualization, Y.D. and C.W.; supervision, C.W.; project administration, C.W.; funding acquisition, C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RGC: The French National Research Agency (ANR)/the Research Grants Council (RGC) Joint Research Scheme (Project number: A_EdUHK802/22; ANR-22-CE54-0014), RGC: PROCORE-France/Hong Kong Joint Research Project (Project number: F-EdUHK803/24, 52627VM), and RGC: General Research Fund (Project number: 18605725). The APC was funded by RGC: The French National Research Agency (ANR)/the Research Grants Council (RGC) Joint Research Scheme (Project number: A_EdUHK802/22; ANR-22-CE54-0014).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Note

¹ “Danjia” is the standard pinyin romanization of the Chinese term for “Tanka”.

References

- Agha, A. (2007). *Language and social relations*. Cambridge University Press.
- Aikhenvald, A. Y. (2007). Grammars in contact: A cross-linguistic perspective. In A. Y. Aikhenvald, & R. M. W. Dixon (Eds.), *Grammars in contact: A cross-linguistic typology* (pp. 1–66). Oxford University Press.
- Aikhenvald, A. Y., & Dixon, R. M. W. (2001). *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford University Press.
- Arppe, A., Lachler, J., Antonsen, L., Trosterud, T., & Junker, M.-O. (2016). Computational modelling of polysynthetic languages. In L. A. Grenoble, & L. J. Whaley (Eds.), *Cambridge handbook of endangered languages* (pp. 104–126). Cambridge University Press.
- Austin, P. K., & Sallabank, J. (Eds.). (2011). *The Cambridge handbook of endangered languages*. Cambridge University Press.
- Bai, Y. (2002). The phonological structure of Xiangzhou Danjia dialect. *Journal of Guangxi Normal University (Philosophy and Social Sciences Edition)*, 38(3), 75–79.
- Bai, Y. (2005). The phonetics of Maocun dialect of Daxu Town, of Lingchuan. *Journal of Guilin Normal College*, 19(2), 23–29.
- Bai, Y. (2006a). On homonyms of Danjia dialect in Xiangzhou County in Guangxi Province. *Journal of Guangxi Normal University (Philosophy and Social Science Edition)*, 42(1), 77–83.
- Bai, Y. (2006b). Slot—The trace of the historical phonetic evolvement of Xiangzhou Danjia dialect. *Journal of Guangxi University for Nationalities (Philosophy and Social Science Edition)*, 28(1), 138–140.
- Bai, Y. (2007a). *A study on the phonetics of the Guangxi Danjia dialect*. Guangxi People's Publishing House.
- Bai, Y. (2007b). Classification of the Danjiahua in southern Guangxi. *Dialect*, 29(3), 274–281.
- Bai, Y. (2007c). The phonetics of Danjia dialect of Yongshui County. *Journal of Guilin Normal College*, 21(3), 1–7.

- Baker, C. (1992). *Attitudes and language*. Multilingual Matters.
- Bapna, A., Wang, T., Bhatia, K., Gu, J., Biderman, S., & Cherry, C. (2022). MASSIVE: A 51-language, 1 million utterance multilingual dataset for language understanding. *arXiv*, arXiv:2204.08582.
- Bird, S. (2022). AI and the world's languages: Automating documentation and revitalisation. In *Proceedings of the 2nd workshop on natural language processing for indigenous languages of the Americas (AmericasNLP 2022)* (pp. 1–13). Association for Computational Linguistics.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press.
- Blust, R. (2006). The linguistic position of the Sama–Bajaw languages. *Oceanic Linguistics*, 45(1), 1–42.
- Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.
- Bradley, D., & Bradley, M. (Eds.). (2002). *Language endangerment and language maintenance*. Routledge.
- CASP. (2018). *CASP qualitative checklist. Critical appraisal skills programme*. Available online: <https://casp-uk.net/> (accessed on 10 March 2025).
- Chan, W. F. (2012). A brief introduction to the phonetics of Tanka dialect in Tai O, Hong Kong. *Journal of Southern Linguistics*, 4, 11–16.
- Chang, S. H., & Zhuang, C. S. (2003). *The dialects of the New Territories*. The Commercial Press.
- Chen, F. T. (2017). *A study on the Danjia dialect in Zhaoping, Hezhou* [Master's thesis, Guangxi Normal University]. China Knowledge Resource Integrated Database.
- Cheng, M. M. (2020). *A study on the phonetics and grammar of the Danjia dialect in Wuzhou* [Master's thesis, Guangxi Normal University]. China Knowledge Resource Integrated Database.
- Cheng, M. M., & Liu, L. Y. (2024). A study on family language use among the Tanka people in Wuzhou. *Chinese National Expo*, 234–236.
- Cheung, S. T. (2020). *Language identity and dialect endangerment—Taking the Tankas in Hong Kong's Cheung Chau as an example* [Bachelor's thesis, The Education University of Hong Kong].
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–476. [CrossRef]
- Enfield, N. J. (2005). Areal linguistics and Mainland Southeast Asia. *Annual Review of Anthropology*, 34, 181–206. [CrossRef]
- Evans, N. (2010). *Dying words: Endangered languages and what they have to tell us*. Wiley-Blackwell.
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters.
- François, A. (2014). Trees, waves and linkages: Models of language diversification. In C. Bowerman, & B. Evans (Eds.), *The Routledge handbook of historical linguistics* (pp. 161–189). Routledge.
- Fung, K. K. (2012a). Phonological characteristics of the Tanka Dialect in Jiusha Village, Dashazhen, Huangpu, Guangzhou. *Journal of Southern Linguistics*, (5), 7–10.
- Fung, K. K. (2012b). The phonetic system and characteristics of Tanka dialect in Shek Pai Wan of Aberdeen of Hong Kong. *Cantonese Studies*, 4–16.
- Fung, K. K. (2015). *The language heritage and cultural transformation of boat people communities in the Pearl River Delta*. Ten Thousand Scrolls Pavilion.
- Fung, K. K. (2023). *Phonological study of the Cantonese sub-dialect in Hong Kong fishing villages*. Ten Thousand Scrolls Pavilion.
- Gal, S., & Irvine, J. T. (1995). The boundaries of languages and disciplines: How ideologies construct difference. *Social Research*, 62(4), 967–1001.
- Giles, H., Bourhis, R. Y., & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307–348). Academic Press.
- Grenoble, L. A., & Whaley, L. J. (2006). *Saving languages: An introduction to language revitalization*. Cambridge University Press.
- Guo, S. H. (2002). *A survey report on the dialect of boat people in Macau* [Master's thesis, Jinan University].
- Haugen, E. (1972). *The ecology of language: Essays by Einar Haugen*. Stanford University Press.
- Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge University Press.
- Higgins, R., & Rewi, P. (Eds.). (2014). *The value of the Māori language: Te hua o te reo Māori*. Huia Publishers.
- Himmelman, N. P. (1998). Documentary and descriptive linguistics. *Linguistics*, 36(1), 161–195. [CrossRef]
- Hornberger, N. H. (2008). *Can schools save Indigenous languages? Policy and practice on four continents*. Palgrave Macmillan.
- Huang, G. F. (2024). *An investigation into the Tanka language of returned overseas Chinese in Poluo Bay Island, Qiaogang Town, Beihai, Guangxi*. World Publishing Corporation.
- Huang, G. F., & Xu, D. D. (2018). The nature of the Danjia dialect of the returned overseas Chinese on Powan Island in Qiaogang Town, Beihai City. *Journal of Hezhou University*, 34(4), 32–38.
- Huang, G. G. (1990). The phonological system and characteristics of the Tanka dialect in Sanya City, Hainan Province. In B. H. Zhan (Ed.), *Proceedings of the second international conference on Cantonese dialects* (p. 131). Jinan University Press.
- Huang, G. G. (1991). The distribution of Chinese dialects in Sanya, Hainan. *Dialects*, 243–254.
- Huang, Z. X. (2015). *A study on the vocabulary of the Danjia dialect in Nanning, Guangxi* [Master's thesis, Guangxi Normal University]. China Knowledge Resource Integrated Database.

- Kiu, K. L. (1984). On some phonetic characteristics of the Cantonese sub-dialect spoken by the boat people from Po Toi Island. *Journal of the International Phonetic Association*, 14(1), 35–37. [CrossRef]
- Ko, S. C. (2021). *A study of the salty water songs of the Tanka boat-dwelling family: Language features* [Bachelor's thesis, The Education University of Hong Kong].
- Leeman, J. (2015). Heritage language education and identity in the United States. *Annual Review of Applied Linguistics*, 35, 100–119. [CrossRef]
- Leonard, W. Y. (2011). Challenging “extinction” through modern Miami language practices. *American Indian Culture and Research Journal*, 35(2), 135–160. [CrossRef]
- Lok, K. H. (2010). *A study of the phonology of Cantonese dialect of Cheung Chau dwellers* [Bachelor's thesis, Hong Kong Shue Yan University].
- Luo, P. S. (2015). *Comparative study of the differences between the Shau Kei Wan waterway dialect and surrounding waterway dialects in Hong Kong* [Bachelor's thesis, Hong Kong Shue Yan University].
- Mao, C. Z. (2015). Review of research on the language of the Dan people. *Theoretical Observation*, 11, 44–49.
- Mao, C. Z. (2019). On language attitudes of the Dan people in Hainan from the perspective of sociolinguistic. *Humanities & Social Sciences Journal of Hainan University*, 37(5), 93–101.
- May, S. (2012). *Language and minority rights: Ethnicity, nationalism and the politics of language*. Routledge.
- McCoy, J. (1965). The dialects of Hong Kong boat people: Kau Sai. *Journal of the Hong Kong Branch of the Royal Asiatic Society*, 5, 46–64.
- Meyer, M., Alam, F., & Darwish, K. (2022). Natural language processing for endangered languages: A survey. *Journal of Artificial Intelligence Research*, 74, 1715–1756.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. [CrossRef]
- Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge University Press.
- Mühlhäusler, P. (1996). *Linguistic ecology: Language change and linguistic imperialism in the Pacific region*. Routledge.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. [CrossRef]
- Pan, J. Y. (2001). The variation of the Raoping Danjia Cantonese dialect and the loss of its communicative function. *Journal of Shantou University (Humanities and Social Sciences Edition)*, 17(3), 36–43.
- Pan, J. Y., & Luo, L. L. (2013). The Danjia people and dialect along the coast of Hailufeng. *Journal of Hanshan Normal University*, 34(2), 99–105.
- Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.
- Pu, J. (2019). *A study on the phonetics and grammar of the boat dialect in the urban area of Guilin* [Master's thesis, Guangxi Normal University]. China Knowledge Resource Integrated Database.
- Schmidt, J. (1872). *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Hermann Böhlau.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23(3–4), 193–229. [CrossRef]
- Stroud, C. (2001). African mother-tongue programmes and the politics of language: Linguistic citizenship versus linguistic human rights. *Journal of Multilingual and Multicultural Development*, 22(4), 339–355. [CrossRef]
- Thomason, S. G. (2001). *Language contact: An introduction*. Edinburgh University Press.
- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- Tong, A., Flemming, K., McInnes, E., Oliver, S., & Craig, J. (2007). Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology*, 12(1), 181. [CrossRef] [PubMed]
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., & Hempel, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. [CrossRef] [PubMed]
- Trudgill, P. (1986). *Dialects in contact*. Blackwell.
- UNESCO. (2003, March 10–12). *Language vitality and endangerment*. International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Fontenay, France.
- United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development* (Resolution adopted by the General Assembly on 25 September 2015, A/RES/70/1). Available online: <https://undocs.org/A/RES/70/1> (accessed on 24 February 2026).
- Wang, C., Peng, D. X. W., Dai, Y. M., & Qi, C. (2025). The endangered Tanka language in Hong Kong: Phonological variations and lexical convergence with Cantonese. *Humanities & Social Sciences Communications*, 12, 1133.
- Ward, J., Besacier, L., Adda, G., & Adda-Decker, M. (2022). How AI can help save endangered languages. *Communications of the ACM*, 65(2), 26–28.

- Wimmer, A., & Glick Schiller, N. (2002). Methodological nationalism and beyond: Nation–state building, migration and the social sciences. *Global Networks*, 2(4), 301–334. [[CrossRef](#)]
- Wu, Y. X. (2007). *A comprehensive analysis of the regional characteristics of the Tai O waterway dialect* [Bachelor's thesis, Hong Kong Shue Yan University].
- Xu, W. H. (2016). A survey of language use of the Raoping Dan people. *Journal of Hanshan Normal University*, 37(2), 47–52.
- Xu, W. H. (2017). *A study on the phonetics of the Raoping Danjia dialect* [Master's thesis, Guangdong Polytechnic Normal University]. China Knowledge Resource Integrated Database.
- Yang, B., & Cheng, M. M. (2020). Literature review of Tanka language in Lingnan area. *Journal of Wuzhou University*, 30(2), 64–70.
- Yang, J. Y. (2019). *A study on the phonetics and vocabulary of the Danjia dialect in Chongkoumen fishing village, Nanlang Town, Zhongshan City* [Master's thesis, Guangxi Normal University]. China Knowledge Resource Integrated Database.
- Yao, Q. Z. (2022). Sound variations of Danjia dialect in Shatian Town, Dongguan City. *Journal of Guangdong Polytechnic Normal University*, (4), 73–79.
- Yeh, E. T. (2007). Cultural and discursive geographies of resistance: Tibetan self-immolation and the spatial politics of protest. *Political Geography*, 56, 1–11.
- Zhang, S., & Huang, X. (1988). Examination of the “Danjia” water people in the Pearl River Estuary. *Social Science Frontline: Ethnology*, 311–324.
- Zhuang, C. S. (2009). Dialects of waterborne residents (Tanka) in Lingnan area. *Cultural Heritage*, 3, 126–132.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.