Unchanging Core Syntax: Gender in Heritage Tamil

MADHUSMITHA VENKATESAN



DEPARTMENT OF HUMANITIES AND SOCIAL SCIENCES

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2026

© Indian Institute of Technology Delhi (IITD), New Delhi, 2026

Unchanging Core Syntax: Gender in Heritage Tamil

by

Madhusmitha Venkatesan

DEPARTMENT OF HUMANITIES AND SOCIAL SCIENCES

Submitted

in fulfillment of the requirements of the degree of

Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

JUNE 2026

Student Research Committee Members

Supervisor:	Prof. Pritha Chandra, Indian Institute of Technology Delhi
Co-Supervisor:	Prof. Roberta D'Alessandro, Utrecht University
Internal Expert:	Prof. Samar Husain, Indian Institute of Technology Delhi
External Expert:	Prof. Rajesh Kumar, Indian Institute of Technology Madras
Chairperson:	Prof. Purnima Singh, Indian Institute of Technology Delhi
External Reviewers	Prof. Maria Polinsky, University of Maryland
	Prof. Artemis Alexiadou, Humboldt-Universität zu Berlin

Declaration

I hereby declare that:

1. The work presented in this thesis is original and has been carried out by me as a student of IIT Delhi.

2. To the best of my knowledge, the work presented in this thesis has not been submitted, either in whole or in part, to any other institution for the award of any degree or diploma.

3. I have adhered to the guidelines prescribed by the Institute for thesis preparation.

4. I have complied with the provisions of the Institute's Honor Code and have neither engaged in nor tolerated any form of academic dishonesty, including plagiarism, intentional data fabrication, data falsification, data misrepresentation, or unauthorized collaboration beyond that permitted by the Student Research Committee, as defined and approved by the Institute.

5. I take full responsibility for all data presented in this thesis and confirm that no data has been collected without appropriate permission. Where required, ethical approval has been obtained for the research, and all relevant approvals and documentation have been included as part of the thesis submission.

6. All sources of information, including data, ideas, and text, have been duly acknowledged and appropriately cited in the thesis.

7. The use of artificial intelligence (AI) tools (if any), including large language models, has been appropriately disclosed. All analysis interpretations, and conclusions presented in

this thesis are my own. Alternatively, I declare that no AI tools, including large language models, have been used in the preparation of this thesis.

8. I understand that any violation of the above provisions may lead to appropriate action as per Institute rules, including withholding of degree approval or revocation of the degree, as applicable under Institute regulations.

Date: 30th June 2026
Place: New Delhi

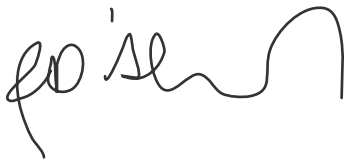
Name of the Student: Madhusmitha Venkatesan
Entry Number.: 2021HUZ8159

Certificate

This is to certify that the thesis – **"Unchanging Core Syntax: Gender in Heritage Tamil"** – submitted by **Madhusmitha Venkatesan** to the Indian Institute of Technology Delhi (IIT Delhi) for the award of the degree of **Doctor of Philosophy** is a record of bona fide research work carried out by her under our supervision and guidance, and it is in conformity with the rules and regulations of IIT Delhi.

She has fulfilled all the requirements for the submission of this thesis which has, in my opinion, reached the standard of fulfilling the requirements for the degree.

To the best of my knowledge, the material contained in this thesis has not been submitted, in part or in full, to any other University or Institute for the award of another degree or diploma.



Pritha Chandra
Professor
Department of Humanities and Social Sciences,
Indian Institute of Technology Delhi,
New Delhi, India
Date:

Roberta D'Alessandro
Professor
Institute for Language Sciences,
Utrecht University,
Utrecht, Netherlands
Date:

Acknowledgements

It is time to write what is perhaps the most notable part of this dissertation: an acknowledgement of everyone who has made an impact on the project. At the top of the list lie my supervisors. I would like to thank Prof. Pritha Chandra for imparting her love for the theory to me, and for teaching me by example to engage with the big questions while not losing focus on the empirical facts. I am profoundly thankful to Prof. Roberta D'Alessandro for all her efforts in making me feel confident in my own work, and for sharing with me her enthusiasm and zest for the theory. Both of them have had a monumental part in shaping the conceptual side of this dissertation.

Alongside my supervisors, I want to thank the other members of my research committee for their help and support throughout: Prof. Purnima Singh for her guidance, Prof. Rajesh Kumar for his technical comments and feedback, and Dr. Samar Husain for kind words and insightful questions. I would also like to thank Prof. Marcel den Dikken and Dr. Silvia Terenghi for their help whenever I was stuck with a theoretical problem. I deeply appreciate the detailed conversations I have had with them over the years. At this juncture I must take a moment to thank all the participants of my studies. They volunteered their time and willingly took part whenever I reached out to them. I sincerely appreciate all their efforts. I would also like to thank Prof. Uma Devi (University of Delhi) and Mr. Velu for their help.

I could not possibly overstate the role played by my peers at MS-621 and MS-615D. Whether in terms of spirited academic debates or frequent coffee breaks, they have been an invaluable support system over the years. Dr. Preeti Kumari, Dr. Chandni Dutta and Dr. Sayantani Banerjee deserve a special mention for all the academic advice they offered to a novice, and for being available every step of the way. I would especially like to thank Preeti and Chandni for the numerous abstracts they have read and critiqued for me, regardless of distance or time zone factors. I also owe an immeasurable debt of gratitude to Biswanath Dash for the multiple sanity

checks since Day One of this PhD. I could not have asked for a better friend, co-worker and critic by my side. I want to extend a heartfelt appreciation to Menka Singh and Therese Liam Tom for all their love and for the brilliant company they provided. I also want to thank Pratikshya Guru, Pranab Bagartti, Avantika Dev, Akhil Abburu, Nayana Raj, Nikhar Agarwal, Kanishka Jain, Greeshma Joseph, Akankshya Misra and Sudipta Saha. All of them believed in me and my work, even on days when I could not. Joaquim Jossy is another friend I want to thank for the weekly debriefs about all things under the sun.

Apart from the academic world, I consider myself fortunate to have a bunch of people who have formed the backbone of my support system. Manal Fathima, Teenu Mathew, and Jyothsna Suresh, with their continued cheerleading from miles away, meant a lot to me. Ms. Chitra Krishnan and Mr. Syriac Mary were constant sources of inspiration, with their ever-expanding love of books and reading. Finally, I have reached the ultimate source of all my strength. I want to thank Amma and Appa for all their support and Maithreyi for her unbridled love. Amma, particularly with her sustained enthusiasm for this project, was a constant source of encouragement for me. I also want to thank Dhananjay for his companionship over the years and for his continued faith in me and this dissertation. I must mention Paati and the entire Gopaladesikan family - particularly Sumedha, Dhanya, Shreemayi, Sriniketh and Sricharan for their curiosity and excitement in my work.

Lastly, I thank the city of Delhi, the indomitability of human spirit, and the theory of generative syntax. All three of them, with their built-in ebbs and flows, have been my guidepost for writing this dissertation.

Madhusmitha Venkatesan

Abstract

Heritage languages are sites of rapid language change, as evidenced by the innovations observed in multiple grammatical domains such as inflectional morphology, case marking, discourse phenomena, etc. However, one domain forms an exception to this by exhibiting stability - core syntax. Ongoing research in the field is focused on defining the bounds of what constitutes core syntax for heritage grammars. This dissertation takes the idea forward by placing core syntax in a grammatical domain that has been established as vulnerable across language acquisition contexts - gender features and gender agreement.

Such a juxtaposition of the reported stability of core syntax and the volatility of gender features in heritage grammars allows this dissertation to explore the ramifications for the structural representation of gender features in heritage grammars. The question is addressed in three ways: (i) Gender assignment to nouns (ii) Dependent marking of gender, or gender agreement (iii) Gender features interacting with structural aspects like possession.

The empirical core of the dissertation is formed by Tamil (Dravidian) acquired as a heritage language in New Delhi. Through a series of picture-description and sentence elicitation tasks, this dissertation examines if heritage Tamil converges to or diverges from the baseline in terms of the structural representations of gender in these domains.

In the domain of gender assignment to nouns, the approach adopted to gender is a configurational one; this dissertation shows that gender features encode representational differences in terms of their location (lexical vs functional heads) and assignment strategies (inherent vs discourse-obtained). Novel empirical evidence from heritage Tamil points towards sensitivity towards these: lexical gender assignment remains stable whereas gender assignment that involves discourse elements undergoes variation in heritage Tamil.

In the domain of gender agreement in adjectives, the dissertation first lays out the structural representation of adjectives, which are also configurational in Tamil. The structure of adjectives requires the application of two mutually exclusive strategies of adjective formation, thus requiring intensive, case-by-case learning. Further, predicative adjectives, to the exclusion of attributive adjectives, participate in agreement relations with nouns. Data from heritage Tamil shows two broad trends: the mechanisms underlying gender agreement on adjectives remain stable, whereas adjective formation shows variation.

The third domain explored in this dissertation is gender and its interaction with possessives. Nouns carrying gender features are disallowed from occurring in the configuration of predicative possession. This constraint is explained as a result of the functional heads involved in possession and their incompatibility with gender features in Tamil. Evidence from heritage Tamil shows that the core structural properties of gender remain stable. The structural constraint is retained in heritage Tamil.

To conclude, this dissertation identifies three grammatical domains where the core-syntactic aspects of gender features remain stable in heritage Tamil. The broad implication of these findings is that we can argue for syntax emerging as a resilient grammatical phenomenon despite the presence of reduced/ restricted primary linguistic data and potential transfer effects from other languages in the environment.

सारांश

हेरिटेज भाषाओँ में व्याकरण का बहुत शीघ्रता से बदलाव देखा जाता है जैसे कि -इन्फ्लेक्शनल मॉर्फोलॉजी, केस मार्किंग, डिस्कोर्स फेनोमेना। इसके विपरीत, एक अनुक्षेत्र है जिसमे स्थिरता देखी जाती है - कोर सिंटैक्स। हेरिटेज भाषाओँ में चल रही रिसर्च इसको परिभाषित करने पर केंद्रित है कि हेरिटेज ग्रामर के लिए कोर सिंटेक्स क्या है। इस शोध प्रबंध में कोर सिंटेक्स को एक ऐसे ग्रामर के क्षेत्र में रखा गया है जिसे भाषा सीखने के संदर्भ में बदलाव के आधीन माना जाता है - जेंडर फीचर्स और जेंडर अग्रीमेंट।

कोर सिंटैक्स की स्थिरता और जेंडर फीचर्स की अस्थिरता के संयोजन का प्रयोग करके यह शोध प्रबंध हेरिटेज ग्रामर में जेंडर फीचर्स के संरचनात्मक प्रतिनिधित्व का परीक्षण करता है। इस प्रश्न को तीन दृष्टिकोण से देखा गया है: (i) संज्ञाओं का जेंडर असाइनमेंट (ii) जेंडर अग्रीमेंट (iii) जेंडर फीचर्स का अधिकार जैसे संरचनात्मक पहलुओं पर परस्पर प्रभाव

इस शोध प्रबंध का अनुभवजन्य भाग नई दिल्ली में हेरिटेज भाषा के रूप में सीखी गई तमिल (द्रविड़) से बना है। चित्रों के वर्णन करने और वाक्य बनाने के माध्यम से, यह शोध प्रबंध जाँच करता है कि क्या हेरिटेज तमिल इन अनुक्षेत्रों में जेंडर के संरचनात्मक प्रतिनिधित्व के मामले में बेसलाइन जैसी है या उससे विपरीत है।

इस शोध प्रबंध में जेंडर असाइनमेंट के लिए कॉन्फ़िगरेशनल एप्रोच अपनाया गया है। यह देखा गया है कि जेंडर फीचर्स अपने स्थान (लेक्सिकल बनाम फंक्शनल हेड्स) और असाइनमेंट रणनीतियों (आंतरिक बनाम डिस्कोर्स से प्राप्त) के सन्दर्भ में पृथक हैं। हेरिटेज तमिल से आते हुए नए अनुभवजन्य प्रमाण इन बातों के प्रति संवेदनशीलता की ओर इशारा करते है कि लेक्सिकल जेंडर असाइनमेंट स्थिर रहता है और डिस्कोर्स तत्वों वाले जेंडर असाइनमेंट में हेरिटेज तमिल में बदलाव आता है।

विशेषण में जेंडर एग्रीमेंट के क्षेत्र में यह शोध प्रबंध पहले विशेषण के संरचनात्मक प्रतिनिधित्व का वर्णन करता है। इसके अतिरिक्त , विशेषणों की संरचना के लिए विशेषण निर्माण की दो रणनीतियों को विषयानुसार सीखने की आवश्यकता होती है। विशेषणात्मक विशेषण में जेंडर एग्रीमेंट नहीं पाया जाता है लेकिन विधेय विशेषणों में एग्रीमेंट पाया जाता है। हेरिटेज तमिल का डेटा दो व्यापक रुझान दिखाता है: विशेषणों पर जेंडर एग्रीमेंट के अंतर्निहित तंत्र स्थिर रहते हैं, जबकि विशेषण निर्माण में भिन्नता दिखाई देती है।

इस शोध प्रबंध में जेंडर फीचर्स और उनका अधिकारवाचक शब्दों के साथ संबंध भी देखा गया है। जिन संज्ञा में जेंडर फीचर्स होते हैं, वे विधेय अधिकार के विन्यास में उपयोग नहीं किये जाते हैं। इस प्रतिबन्ध को अधिकार में शामिल फंक्शनल हेड्स और तमिल में जेंडर फीचर्स के साथ उनकी असंगति के रूप में समझाया गया है। हेरिटेज तमिल के डेटा से पता चलता है कि जेंडर की मुख्य स्ट्रक्चरल प्रॉपर्टीज़ स्थिर रहती हैं।

निष्कर्ष के तौर पर, यह शोध प्रबंध तीन ग्रामेटिकल डोमेन की पहचान कराता जिसमे हेरिटेज तमिल में जेंडर फीचर्स के मुख्य सिंटैक्टिक पहलू स्थिर रहते हैं। इसके आधार पर हम यह तर्क दे सकते हैं कि भाषा अधिग्रहण के प्रसंग में सीमित प्राइमरी लिंगविस्टिक डेटा और दूसरी भाषाओं से संभावित ट्रांसफर प्रभावों के होते हुए भी सिंटैक्स सुदृढ़ रहता है।

Contents

Certificate i

Acknowledgements ii

Abstract iv

Hindi Abstract vii

List of Figures xi

List of Tables xii

Abbreviations xiii

1 Introduction 1

1.1 Introduction 1

1.2 About Heritage Languages 4

1.3 Gender and Gender Agreement 5

1.4 Gender in Heritage Languages 7

1.5 The Empirical Core of the Dissertation 10

1.5.1 Gender and Gender Agreement in Tamil 11

1.5.2 Gender and Gender Agreement in Hindi-Urdu 14

1.5.3 Introducing Heritage Tamil in New Delhi 16

1.5.4 Methodology 17

1.6 Core Research Questions 17

1.6.1 Dissertation Outline 18

2 Gender in Nouns **21**

2.1 Introduction . . . 21

2.2 The Gender Systems of Tamil and Hindi-Urdu . . . 22

2.2.1 Gender in Tamil - Empirical Facts . . . 22

2.2.2 Gender in Hindi-Urdu - Empirical Facts . . . 23

2.2.3 Generalisations . . . 25

2.3 The Role of Structure in Gender Assignment . . . 25

2.3.1 Picallo and Longobardi's GenP System . . . 26

2.3.2 Alexiadou's Bipartite Distinction . . . 29

2.3.3 Kramer's Distributed Morphology-Based Account . . . 31

2.3.4 Sigurðsson's Edge Computation Account . . . 33

2.3.5 Interim Summary I . . . 36

2.4 Structural Representations of Gender . . . 37

2.4.1 Representation of Gender in Tamil . . . 41

2.4.2 Representation of Gender in Hindi-Urdu . . . 45

2.4.3 Interim Summary II . . . 48

2.5 Current Proposal: Implications for Language Contact . . . 49

2.5.1 What do existing accounts tell us? . . . 49

2.5.2 The Current Study . . . 53

2.5.2.1 Questionnaire and Methodology . . . 54

2.5.2.2 Hypotheses and Research Questions . . . 57

2.5.2.3 Results of the Study . . . 57

2.5.2.4 Discussion . . . 62

2.6 Conclusion . . . 64

3 Gender in Adjectives **66**

3.1 Introduction . . . 66

3.2 Adjectives in Tamil . . . 67

3.2.1 Existing Accounts: There are no adjectives in Tamil . . . 67

3.2.2 The Current Proposal: Distinct Adjectives in the Syntax . . . 71

3.3 Deriving Adjectives in Tamil . . . 74

3.3.1 A Two-way Distinction . . . 74
3.3.2 Deriving the Two-way Distinction . . . 77
3.3.3 Attributive Adjectives . . . 79
3.3.4 Predicative Adjectives and Gender Agreement . . . 82
3.3.4.1 Two Different Agreeing Heads . . . 89
3.3.5 Interim Summary . . . 91
3.4 Adjectives in Hindi-Urdu . . . 92
3.5 Current Proposal: Implications for Language Contact . . . 95
3.5.1 Existing Accounts of Adjectives in Heritage Grammars . . . 95
3.5.2 The Current Study . . . 98
3.5.2.1 Questionnaire and Methodology . . . 99
3.5.2.2 Hypotheses and Research Questions . . . 102
3.5.2.3 Results of the Study . . . 105
3.5.2.4 Discussion . . . 111
3.6 Conclusion . . . 115

4 Gender in Possessives **117**
4.1 Introduction . . . 117
4.2 Why Possessives? - The Gender Connection . . . 118
4.2.1 Possessives in Tamil - Empirical Facts . . . 120
4.2.2 Possessives in Hindi-Urdu - Empirical Facts . . . 123
4.2.3 Generalisations . . . 124
4.3 Structural Representation of Possessives . . . 125
4.3.1 Attributive Possession . . . 125
4.3.2 Predicative Possession . . . 133
4.3.2.1 Biological Gender, *n*, and Predicative Possession . . . 139
4.3.2.2 Semantic Constraint on Possession? . . . 141
4.3.3 Interim Summary . . . 143
4.4 Current Proposal: Implications for Language Contact . . . 143
4.4.1 Existing Accounts of Possession in Heritage Grammars . . . 143
4.4.2 The Current Study . . . 146

4.4.2.1 Questionnaire and Methodology 147
4.4.2.2 Hypotheses and Research Questions 151
4.4.2.3 Results of the Study 153
4.4.2.4 Discussion 156
4.5 Conclusion 160

5 Conclusion **161**
5.1 Summary of the Dissertation 161
5.2 Key Takeaways 164
5.2.1 Absence of Cross-Linguistic Transfer 164
5.2.2 The Resilience of Core Syntax in Heritage Languages 166
5.3 Suggestions for the Way Forward 168
5.3.1 More Dimensions of Core Syntax 168
5.3.2 More Empirical Domains 169
5.4 General Conclusions 171

Bibliography **172**

Bio Data **181**

List of Figures

1.1 Map of India 11

2.1 Prompts for Lexically Specified Gender 54
2.2 Prompts for Pronouns 55
2.3 Prompts for Common Nouns 56
2.4 Distribution of MASC and FEM Pronouns 60
2.5 Distribution of MASC and FEM Common Nouns 62

3.1 Prompts for Attributive Adjectives 100
3.2 Prompts for Predicative Adjectives 101
3.3 Distribution of Adjectives According to their Formation 106
3.4 Distribution of Attributive and Predicative Adjectives 109

4.1 Prompts for Attributive Possession 148
4.2 Prompts for Predicative Possession: Non-Gendered Nouns 150
4.3 Prompts for Predicative Possession: Gendered Nouns 151

List of Tables

2.1 Gender in Tamil and Hindi-Urdu 25
2.2 Results of the study 62

3.1 Adjective Formation in Tamil 92
3.2 Elicited Adjectives According to their Formation 106
3.3 Distribution of Attributive and Predicative Adjectives- Absolute Values 109

4.1 Distribution of Attributive and Predicative Possession- Absolute Values 155

List of Abbreviations

ACC	Accusative Case	INTF	Intensifier
ADJR	Adjectivizer	LOC	Locative Case
CLI	Cross-Linguistic Influence	MSG	Masculine Singular
COP	Copula	NEG	Negation
CPM	Conjunctive Participial Marker	NMLZ	Nominalizer
DAT	Dative Case	NOM	Nominative Case
ERG	Ergative Case	PERF	Perfective Aspect
FSG	Feminine Singular	PL	Plural
FUT	Future	PLD	Primary Linguistic Data
GEN	Genitive Case	POSS	Possessive
HL	Heritage Language	PST	Past
INF	Infinitive	REL	Relativizer