Gender Agreement in a Tamil-Hindi Bilingual Situation: The Role of Feature Valuation

MADHUSMITHA VENKATESAN, *Indian Institute of Technology Delhi*

ABSTRACT

Within the Generative paradigm, variation is understood in terms of features. A crucial mechanism for this is the Borer-Chomsky Conjecture, explained in Baker (2008) as: 'All parameters of variation are attributable to differences in the features of particular items in the lexicon.' This paper is an attempt to understand a case of asymmetric bilingualism of Hindi-Urdu (Indo-Aryan) and Tamil (Dravidian) and to explain the resultant changes in gender agreement in terms of featural configurations of functional heads. The empirical core is set in New Delhi, with simultaneous bilinguals acquiring Tamil (L1) at home and Hindi-Urdu (L2) from the external environment. While both grammars mark gender information on the verb, Hindi-Urdu has grammatical and biological gender. In Tamil, on the other hand, gender is purely semantic. The Hindi-Urdu grammar of this bilingual population appears to find gender agreement challenging.

This paper adopts a representational approach; the loss of Gender is analysed as the deletion of an uninterpretable valued feature on a functional head. The first approach is to posit the uninterpretability of the feature as the cause of deletion. The inability of this claim to hold up empirically is then taken to mean that the explanation for the deletion of the feature lies in its other property: Value. Reanalysis of the change in contact situations reveals that losing valued features could simply be a strategy adopted by languages in an effort to be more parsimonious.

1 Introduction

Contact situations are one of the most significant contributors to language variation and change. Such changes are crucial in enriching our existing knowledge on the shape of Universal Grammar, especially with regard to its extent of plasticity. Within the generative paradigm, understanding variation in terms of features has been a rewarding exercise as it affords us a microscopic view of the way in which languages undergo change. This paper attempts to contribute to the discussion by introducing the factor of a contact situation. The objective of this paper is to understand the features underlying gender agreement in Hindi-Urdu when placed in the bilingual context of simultaneous bilinguals of Hindi-Urdu and Tamil.

1.1 Gender as a Category

Gender is primarily a system of nominal categorization; it is a way of organising nouns into different categories based on some inherent attributes that the items may possess (Corbett, 1991). There are multiple aspects of gender which make its study essential; the acquisition of Gender is intricately connected to structure, the presence of gender in the world's languages is highly varied, and its omnipresent involvement in agreement relations reveals much about the role of features in agreement operations. All this combined offers an insightful window into the mental architecture of Language.

Gender in languages is broadly divided into two kinds: Natural/Biological or Semantic gender (BG), and Grammatical Gender (GG). The former is a system in which meaning plays a central role. Only animate nouns are allotted a gender, and their gender value correlates with the biological sex of the referent. Thus, nouns denoting a male human or animal are assigned [MASCULINE] gender, and nouns denoting female humans or animals are assigned [FEMININE] gender. Crucially, these are nouns where the gender value of the noun is grounded in its meaning. Grammatical Gender, on the other hand, should be understood as a more complex and abstract concept than just the co-referent of biological sex. GG may or may not be congruent with natural/biological gender, animacy or other related semantic properties. Grammatical Gender is often referred to as arbitrary gender (Kramer, 2015), as neither is there consistency in the assignment of gender to inanimate objects nor do the intrinsic properties of the noun have any role to play in its gender assignment.

1.2 Gender and Language Contact

Whenever the context is bilingual, grammatical gender seems to be vulnerable. Loss of gender is cross-linguistically common in language contact situations. This is well recorded in the literature (Sánchez et al., 2022; Igartua, 2019; Lohndal & Westergaard, 2016; Kramer, 2014; Karatsareas, 2009; Oliphant, 1998; Gumperz & Wilson, 1971). These are all cases of typologically unrelated languages losing grammatical gender, and the reason has canonically been attributed to sustained contact with a gender-less language. This brief but telling literature survey forms the basis for this study: Empirical evidence points towards gender being vulnerable in different kinds of language contact situations. What could possibly motivate such a tendency in gender agreement? How can we capture this using the theoretical mechanisms of generative grammar?

1.3 Gender in the Generative Theory: An Uninterpretable and Valued Feature

Features have been central to generative theory almost from its inception (Chomsky, 1965). Since then, features have evolved as crucial for operations such as MERGE, MOVE and AGREE; it is the features of a lexical item that decide which other lexical items it may combine with, and features are integral for forming syntactic dependencies between a Probe and a Goal, leading to agreement. This naturally leads to the next question: How do we represent gender as a feature? According to the Universal DP structure (1), as proposed

by Polinsky (2016) and Carstens (2000), each projection inside the DP is associated with different phi-features. Person features are assumed to be a property of the determiner D, number features are contained in the Num head, and gender features are considered a lexical property of the noun itself.

(1)
```
              DP
             /  \
          D      NumP
       Person   /    \
             Num      NP
            Number   Gender
```
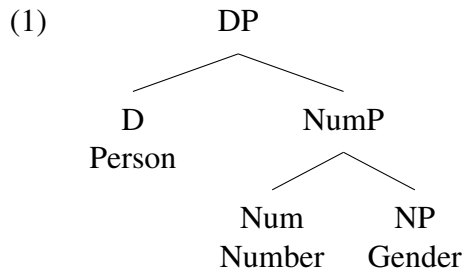
Sigurdsson (2019) proposes an analysis wherein the interpretable but unvalued Gender feature on D acts as a Probe and looks for a value. It finds its Goal in the uninterpretable but valued Gender feature on the NP. A relation of Agree is established between the two and this is how D gets its valuation for Gender. Such a relation with the Interpretable feature being the Probe and the Uninterpretable feature being the Goal is not permitted in the Chomskyan system (Chomsky, 2000). However, Pesetsky & Torrego (2007) provides a framework within which the interpretable status of a feature does not prevent it from being a Probe; as long as a feature requires valuation through syntax, it can act as a Probe.

When we approach Gender as a feature on a functional head, it opens up the possibility of analysing the changes taking place in the contact situation from the perspective of the Borer-Chomsky Conjecture. The Borer-Chomsky Conjecture given by Borer (1984), Chomsky (2001) and explained in Baker (2008) states that all parameters of variation are attributable to differences in the features of particular items (eg, functional heads) in the lexicon. This approach affords us a microscopic view of how languages show variation and will have interesting implications when extended to contact situations, which differ from diachronic change in having to factor in external forces such as the influence of another language.

Following the Borer-Chomsky Hypothesis which considers even large-scale language change as an alteration of the featural composition of functional heads, the loss of grammatical gender in a language can be characterised as the loss of the uninterpretable [uGen:FEM] feature on n.

2 The Current Problem

The empirical core of this study is formed by the grammar of Tamil-Hindi/Urdu bilinguals residing in New Delhi. Here we present a brief overview of the gender systems of the two grammars.

2.1 Gender in Hindi-Urdu

Hindi/Urdu is an Indo-Aryan language with semantic as well as formal systems of gender Corbett (2013). All nouns, regardless of their status with respect to animacy, are assigned a gender value. There are two gender categories in Hindi, Masculine and Feminine, and every noun in the language is allotted one of the two. Additionally, Hindi/Urdu also has a rich agreement paradigm; gender (and other phi features) are marked on a range of functional heads: D (2a , 2b) , ADJ (2c , 2d), T (2e , 2f) and *v* (2g , 2h) .

(2)   a. mer-ii kitaab
           my-F book.F
           'My book'

       b. mer-aa betaa
           my-M son.M
           'My son'

       c. lamb-ii ladki
           tall-F   girl.F
           'Tall girl'

       d. baD-aa ghar
           big-M house.M
           'Big house'

       e. wah ladkaa roTii   khaat-aa   hai
           DEM boy.M bread.F eat.PRS-3MS be.PRS
           'That boy eats bread'

       f. wah ladkii ghar   jaa   rah-ii   th-ii
           DEM girl.F house.M go.INF PROG-3FS be.PST-3FS
           'That girl was going home'

       g. vijay-ne   roTii   khaa-yii
           Vijay.M-ERG bread.F eat.PST-3FS/*3MS
           'Vijay ate (the) bread'

       h. vijay-ne   roTii khaan-ii   chaah-ii
           vijay.M-ERG bread.F eat.INF-3FS want-PST-3FS
           'Vijay wanted to eat bread'

2.2 Gender in Tamil

Tamil is a Dravidian language with a gender system that is entirely semantic , i.e., based on the natural gender of the noun's referent. Consequently, there exists a sharp divide between [+HUMAN] and [-HUMAN] nouns. Thus, all nouns denoting human beings are encoded with a gender value that corresponds to the biological sex of the referent (MASC (3a) or FEM (3b)), and nouns denoting animate non-human (3c), as well as inanimate entities (3d), exhibit the NEUTER or non-gendered marker. The gender value of the noun then manifests

in verbal morphology, as seen in (3). Agreement in Tamil is displayed only on verbs. No other functional heads such as adjectives, determiners, etc participate in agreement.

(3) a. anda paiyan va-nd-aan
         DEM boy.M come-PST-3MS
         'That boy came'

     b. anda ponnu va-nd-aa
         DEM girl.F come-PST-3FS
         'That girl came'

     c. maadu va-nd-udu
         cow.N come-PST-3NS
         '(a/the) cow came'

     d. seidi   va-nd-udu
         news.N come-PST-3NS
         '(the) news reached (us)'

As we can see from the examples above, there are two key differences between the gender systems of (L2) Hindi-Urdu and (L1) Tamil:

- Hindi-Urdu has grammatical as well as biological gender; Tamil has only biological gender.

- Hindi-Urdu has multiple agreeing functional heads- determiners, adjectives, verbs, whereas only verbs agree in Tamil.

2.3 The Target Population

The target population comprises early childhood bilinguals who have spent all or most of their lives in Tamil-speaking families in New Delhi. The rationale behind selecting this group is that by growing up in a bilingual environment (Tamil at home and Hindi/Urdu in the larger society), they would have had access to input from both languages during their initial stage of language acquisition, thus making them early-childhood bilinguals. These speakers could be simultaneous or sequential bilinguals. A

Data on L2 Hindi/Urdu was collected from L1 speakers of Tamil who have grown up in New Delhi. Biological gender mismatches were found in abundance in the data, as seen in (4):

(4) a. *mer-aa maa
         my-M   mother.F
         'My mother'

     b. *usk-aa   betii
         his/her-M daughter.F
         'His/her daughter'

c. *acch-aa ladkii
   good-M girl.F
   'Good Girl'

d. *ach-aa kitaab
   good-M book.F
   'Good book'

e. *sast-aa ghaDii
   cheap-M watch.F
   'Cheap watch'

The examples in (4) demonstrate the absence of gender agreement within the DP, despite the gender value of the nouns 'mother', 'daughter' and 'girl' being readily available to the speakers via external context. We also observe this effect indiscriminately on biological (4a, 4b, 4c) and grammatical (4d, 4e) gender-marked nouns.

The question to ask here is: Why is gender agreement so vulnerable in the grammar of these simultaneous bilinguals? What underlying mechanism of gender agreement is causing this effect?

3 Analysis

This section proceeds to address the question of why gender may be affected so severely in contact situations. The approach adopted here is layered: It is well established (Kramer, 2014, 2015) that gender as a feature is uninterpretable and valued. Our first step would be to posit the uninterpretability of the feature as the cause for its vulnerability. This claim is then held up against cross-linguistic empirical evidence, which then leads us to assess the other attribute of the feature: its value.

3.1 On the Vulnerability of Interpretable Features

Features are deemed uninterpretable when they cannot be meaningfully interpreted at the C-I component or the LF interface. Such features are also considered to be more vulnerable and likely to be lost when a language undergoes change. Van Gelderen (2019) provides an account where a reanalysis of contact-induced change reveals that it is indeed the uninterpretable features that are lost. Supporting evidence for this claim comes from two phenomena: Negation Concord and Pro-Drop, both of which tend to be discarded from a grammar when it enters into contact with another language.

In certain varieties of English, a clause contains more than one negation element, but only one of them is interpretable and meaningful (5). The other negation element is considered to be superfluous and can be reinterpreted as an uninterpretable NEG feature [uNEG].

(5) When my granddaddy dead, I ain't had no children
     When my grandfather died, I did not have any children yet.

It is reported that Negation Concord is not found in contact languages. Van Gelderen (2019) analyses the inability of Negation Concord to be retained in a contact situation as an instance of the loss of the [uNEG] feature in contact. The lexical item with the second [NEG] feature is deleted because the [NEG] here is uninterpretable and thus, vulnerable.

The second source of examples comes from Pro-Drop in language contact. Pro-Drop or Null Subject is traditionally considered to be a category-defining structural parameter. But in terms of features, it can also be described as an uninterpretable feature [uD] on T. Adger (2003) states that the [uD] feature on T is a strong feature, a requirement for a DP that drives movement of the structurally most appropriate DP to [Spec, TP]. This feature on T, too, is uninterpretable. Van Gelderen's analysis of the loss of Pro-Drop in contact situations draws upon this fact: Pro-Drop does not feature in contact situations because the feature controlling it is uninterpretable and hence vulnerable.

The interim conclusion is that all uninterpretable features are vulnerable and therefore must be let go of in contact situations, i.e., situations which force languages to change. The next section takes this idea further, and we immediately encounter a challenge while looking at another uninterpretable feature in action: Agreement controlled by T.

3.2   Not all Uninterpretable Features are Vulnerable

Negation Concord and Pro-Drop in contact situations are certainly indicative of an analysis that banks on the vulnerability of uninterpretable features in general. However, such an account falls short when faced with an obstacle in the form of the other uninterpretable features on T. T is the locus of a bundle of phi-features. The phi features on T, which are responsible for agreement on the tensed verb, are all uninterpretable. According to the idea developed by Van Gelderen, they must all be vulnerable. What we see in contact situations, on the contrary, is that this prediction does not bear out.

Subject-Verb agreement (controlled by T) is not reported as a vulnerable phenomenon in contact situations. It is resilient despite being operated by uninterpretable features. This paves the way for a revision of the claim that uninterpretable features are vulnerable in contact. In the next section, we therefore proceed to look closely at the other defining property of gender features: Inherent Value.

3.3   The Role of Valued Features

Let us revisit all the features discussed so far: [uD] (Pro-Drop), [uNEG] (Negation Concord), [uPhi:] (T Agreement) and [uGen:+FEM] (Gender). Van Gelderen (2019) presented the argument that uninterpretable features are vulnerable in language contact situations and therefore are lost in those contexts. The loss of grammatical gender could be analysed in the same manner. However, such an account fails to explain why T-Agreement, which is also operated by uninterpretable features, is not affected in contact situations.

We then turn to the other property of all these features: Value. Out of the four features mentioned above, three are deleted in language contact situations (Negation Concord, Pro-Drop and Gender Agreement), and one is not (T-Agreement). While uninterpretability is

not helpful in explaining this divide, Valuation is. Upon closer inspection we find a thread of commonality across those features that do get deleted: either they are inherently valued [uGen:+FEM], or they are features that do not require any value [uNEG], [uD]. Both types imply that these features do not need to participate in any structural relation such as Agree in order to get a value. The feature that does not get deleted (T-Agreement), on the other hand, is one that requires a value and will Probe the derivation in search of a Goal. This marks a crucial difference between the two types of features: a. Features that are not at all in need of Valuation. b. Features that will look for Valuation.

Gender Features belong to the former category. By virtue of being inherently valued, gender features simply do not need to act as probes. Since such features are not going to participate in structural relations, they have no function to perform in tying the structure together. Hence, they can be let go of, and languages do end up deleting them in contact situations, especially when there is a catalyst in the form of another language that does not have this feature to begin with.

This paper makes the suggestion that the grammar of Hindi-Urdu, in contact with Tamil, is willing to let go of its inherently valued gender features in an effort to economise its own system

4 Conclusion

The central prediction made by this paper was that valued features are dispensable in a contact situation, as they are not useful for the structure. This goes on to strengthen the claim that the behaviour of languages in a contact situation is no different from a system in isolation, at least with respect to the core property of economy. This claim needs to be explored further, both empirically, as well as in terms of the conceptual implications of value in a feature.

References

Adger, David. 2003. *Core syntax: A minimalist approach*, vol. 20. Oxford University Press Oxford.

Baker, Mark. 2008. The macroparameter in a microparametric world. *The Limits of Syntactic Variation* 351–374.

Borer, Hagit. 1984. Restrictive relatives in modern hebrew. *Natural Language & Linguistic Theory* 2(2). 219–260.

Carstens, Vicki. 2000. Concord in minimalist theory. *Linguistic inquiry* 31(2). 319–355.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. MIT press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework (mitopl 15). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* 89–155.

Chomsky, Noam. 2001. Derivation by phase. *Ken Hale: A Life in Language* 1–52.

Corbett, Greville G. 1991. *Gender*. Cambridge University Press.

Corbett, Greville G. 2013. Systems of gender assignment (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*, Zenodo. doi:10.5281/zenodo.7385533. https://doi.org/10.5281/zenodo.7385533.

Gumperz, John J & Robert Wilson. 1971. Convergence and creolization. *Pidginization and creolization of languages* 151. 167.

Igartua, Iván. 2019. Loss of grammatical gender and language contact. *Diachronica* 36(2). 181–221.

Karatsareas, Petros. 2009. The loss of grammatical gender in cappadocian greek. *Transactions of the Philological Society* 107(2). 196–230.

Kramer, Ruth. 2014. Gender in amharic: A morphosyntactic approach to natural and grammatical gender. *Language sciences* 43. 102–115.

Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford University Press.

Lohndal, Terje & Marit Westergaard. 2016. Grammatical gender in american norwegian heritage language: Stability or attrition? *Frontiers in psychology* 7. 344.

Oliphant, Katrina. 1998. Acquisition of grammatical gender in italian as a foreign language. *Canadian modern language review* 54(2). 239–262.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture* 262–294.

Polinsky, Maria. 2016. Agreement in archi from a minimalist perspective. *Archi: Complexities of agreement in cross-theoretical perspective* 184–232.

Sánchez, Liliana, José Camacho, Elisabeth Mayer & Carolina Rodríguez Alzza. 2022. Gender agreement in a language contact situation. *Languages* 7(2). 81.

Sigurdsson, Halldór Ármann. 2019. Gender at the edge. *Linguistic Inquiry* 50(4). 723–750.

Van Gelderen, Elly. 2019. Are uninterpretable features vulnerable? *Theoretical Linguistics* 45(3-4). 275–282.