DE GRUYTER Open Linguistics 2025; 11(1): 20250073

Research Article

Alexander Angsongna*

Phonotactic constraints and learnability: analyzing Dagaare vowel harmony with tier-based strictly local (TSL) grammar

https://doi.org/10.1515/opli-2025-0073
Received November 8, 2024; accepted November 13, 2025; published online December 12, 2025

Abstract: This paper examines vowel harmony in Dagaare using the Tier-Based Strictly Local (TSL) framework, focusing on tongue root, rounding, backness, and height harmonies. While vowel harmony in Niger-Congo languages, particularly Dagaare, has been explored from phonetic, phonological and typological perspectives, computational insights remain limited. The study applies the TSL framework to model the phonotactic constraints governing harmonic patterns, projecting only harmony-relevant features onto a tier to capture non-local dependencies while ignoring irrelevant segments. This approach allows for precise modeling of agreement among non-adjacent vowels and demonstrates that all harmony types in Dagaare can be represented within a single TSL grammar, avoiding the need for separate tiers for each feature. The findings indicate that the Dagaare system is robustly learnable from surface data under TSL constraints, offering a computationally tractable path for both human and machine learners. A key limitation is also identified: TSL fails to account for harmony exceptions in morphologically complex words, such as compounds, due to its lack of morphological domain sensitivity. The study contributes to the typological understanding of Dagaare, illustrates the utility of TSL for modeling complex harmony systems, and recommends extensions such as domain-sensitive tier projection to better handle morphologically complex contexts.

Keywords: Dagaare; vowel harmony; tier-based strictly local; phonotactic constraints; learnability

1 Introduction

Vowel harmony is a widely studied phenomenon that occurs in many languages, each with distinct phonological properties and varying levels of productivity. This phonological process, whereby vowels within a certain domain harmonize in specific features, is prevalent in the Niger-Congo languages (Boyd 2024; Casali 2003, 2008, 2016), Nilo-Saharan languages (Rose 2018), and Chadic languages (Pearce and Lovestrand 2024). Outside Africa, vowel harmony is observed across diverse language families worldwide (Archangeli and Yip 2019; Casali 2008), as surveyed extensively in Ritter and van der Hulst (2024).

Vowel harmony operates within several phonological dimensions, including the advancement and retraction of the tongue root (ATR/RTR harmony), front/back vowel positioning (palatal harmony), and height and rounding (labial harmony) (Bakovic 2000, 2003). For instance, ATR harmony is particularly common in the Niger-Congo and Nilo-Saharan families (Casali 2008), while labial harmony is seen in African languages like Nawuri (Casali 1995b), Dagaare (Angsongna 2023; Bodomo 1997), the Fante dialect of Akan (Dolphyne 1988; O'Keefe 2004), and some

***Corresponding author: Alexander Angsongna**, Department of African Studies (Linguistics & Literatures), University of Vienna, Spitalgasse 2, Court 5, A-1090 Vienna, Austria, E-mail: alexander.angsongna@univie.ac.at. https://orcid.org/0000-0002-0978-8039

Mbam languages (Boyd 2015). Height harmony, on the other hand, is more frequent in the Chadic language family (Pearce and Lovestrand 2024).

In Dagaare, vowel harmony has been documented in studies, focusing on ATR and rounding harmony (Bodomo 1997), with Optimality Theory analyses (Angsongna 2023) and studies into the acoustic dimensions of ATR harmony (Ozburn et al. 2018). The current paper presents a computational analysis of Dagaare vowel harmony using a Tier-based Strictly Local (TSL) approach (Aksënova et al. 2024; Clements 1985; Heinz 2011, 2018; Heinz et al. 2011). This approach provides a systematic, learnable model to capture the phonotactic constraints underlying harmony patterns in Dagaare, advancing both theoretical and empirical understandings of vowel harmony.

2 Language

Dagaare (ISO-639-3) is part of the Mabia[1] group within the Niger-Congo language family (Bodomo 2020, 1994). It is primarily spoken in northwestern Ghana, including the Upper West Region and parts of the Savannah Region, as well as in areas of Burkina Faso and Ivory Coast, with approximately two million speakers. Closely related varieties include [dàgàrà =(Dagara)] which includes the Lobr/Dagara-Lobr variety spoken around Nandom and Dissin in Burkina Faso, [wáálí (=Waale)], and [bìrìfɔ̀ = Berefɔ)], forming a continuum of mutual intelligibility to varying degrees (Bodomo 1997). Within the broader Mabia group, Dagaare is related to languages such as Gurenɛ (Farefare), Safaliba, Moore, Dagbani, Kusaal, Mampruli, Buli, and Konkomba.

According to Bodomo (1997), Dakubu (2005), and Angsongna and Akinbo (2022), Dagaare comprises four main regional varieties: Northern Dagaare, Central Dagaare, Southern Dagaare, and Western Dagaare, each exhibiting internal variation. Figure 1 presents the four geographically distributed varieties of the language.

Further studies, such as those by Mwinlaaru (2021) and Mwinlaaru and Yap (2017, 2021), identify six principal dialects: Central Dagaare, Lobr, Wule, Southern Birifor,[2] Northern/Malba Birifor, and Waali. A more recent study by Angsongna and Bodomo (forthcoming) reports up to eight main dialects, reflecting the linguistic richness and internal diversity of the language. However, as this study focuses specifically on vowel harmony phenomena, it does not aim to establish a definitive dialectal classification of Dagaare. Therefore, no further discussion of dialectal divisions is provided.

For the purposes of this research, data are drawn from mid-central Dagaare, specifically the variety spoken in Sombo, a community in the Nadowli-Kaleo District of Ghana's Upper West Region.

2.1 Data sources and methods

The primary data for this study originates from the Central Dagaare variety spoken in the Nadowli Kaleo district of Ghana, specifically in the Sombo area. Data collection took place in Ghana from March to April 2018, involving a total of twenty-three speakers. However, most data was gathered from ten key speakers – seven males and three females. The elicitation process included wordlists, phrases, and sentences based on the SIL Comparative African Wordlist (SILCAWL) (Snider and Roberts 2004), as well as short stories, songs, and descriptions of local events and cultural practices. Additional data from existing literature supplemented these findings. As a native speaker, I also contributed supplementary data using my own linguistic intuitions and judgments.

1 For more information about Mabia language group, its etymology, geography and genetic features, see Bodomo (2020). There is the possibility of the existence of a fifth variety of Dagaare, Eastern Dagaare (see Angsongna and Bodomo forthcoming).

2 The terms *Birifor* and *Berefɔ* are used interchangeably to refer to both the language and its speakers. *Berefɔ* is the standard orthographic form, while its phonetic realization is [bìrìfɔ̀]. The form *Birifor*, which also appears in written contexts, appears to be an anglicized rendering of the name.

Figure 1: Geographic distribution of varieties of Dagaare (Angsongna and Akinbo 2022).

For data recording, a Shure WH30XLR cardioid condenser headset microphone and a Rode NGT2 super-cardioid condenser shotgun microphone was used, both at a 48 kHz sampling rate and 16 bit depth and connected to a Zoom Q8 camera.

All data in this study are in broad phonetic transcription.[3]

2.2 Vowel inventory structure

Asymmetries in tongue root harmony have been widely examined in descriptive, typological, and theoretical research. Some languages, such as Diola-Fogny, exhibit a tendency for retracted vowels to assimilate to advanced vowels, while others, like Yoruba, favor spreading the retracted tongue root feature (Archangeli and Pulleyblank 1989). Casali (2024) notes that the dominant tongue root feature correlates with the vowel inventory structure of a language. Languages that contrast tongue root position in high vowels are classified as /2IU/, while those with contrast only in non-high vowels are labeled /1IU/ (Casali 2008, 2024).

The /2IU/ group includes fully symmetric systems like Degema, which has ten vowels equally divided between ATR and RTR pairs, and asymmetric systems like Akan (nine vowels) and Kinande (seven vowels). In the /1IU/ category, Yoruba exemplifies this structure.

3 For the sake of clarity, as reflected in the sound inventory, the following orthographic forms are represented phonetically as follows: y [j], ky [ʧ], gy [ʤ], ny [ɲ], ng [ŋ], and ngm [ŋm].

Dagaare was initially described as having a nine-vowel asymmetric system (Bodomo 1997) with four pairs of contrastive non-low vowels and a non-contrastive low vowel [a]. However, recent findings (Angsongna and Akinbo 2022; Ozburn et al. 2018) indicate that Dagaare has a fully symmetric ten-vowel system, with balanced ATR and RTR vowels, as shown in (1).

(1) Ten-vowel /2IU/ inventory of Dagaare

High	/i/		/u/
	/ɪ/		/ʊ/
Mid	/e/		/o/
	/ɛ/		/ɔ/
Low		/ə/	
		/a/	

Under this system harmony is realized in two ways: (i) within a room morpheme, all vowels must generally come from a set of vowels sharing same tongue root feature. This means that advanced and retracted vowels do not occur together root-internally and (ii) affix vowels alternate in agreement with vowels of the root to which they are attached. The examples below are consistent with the above generalizations.

(2) Root-internal tongue root harmony[4]

ATR	Gloss	RTR	Gloss
bíé	'child.SG'	píɛ́	'roof.SG'
dúŋ-ó	'animal-SG'	dʊ́ŋ-ɔ́	'mosquito-SG'
nə̀ɲìg-é	'thief-SG'	zɔ̀mìŋ-ɛ́	'friend-SG'

As shown in the above data, harmony is exhibited in roots with more than one vowel while suffixes symmetrically harmonize with vowels of the roots. This is referred to as root-/stem-controlled harmony (Bakovic 2003, 2000; Clements 1985). Such symmetric patterns according to Casali (2003) correlate with the general assumption that /2IU/ languages have the tendency of exhibiting [ATR] dominance – a situation in which an [RTR] suffix shows up as [ATR] because of an ATR root.

3 Vowel harmony types in Dagaare

A number of vowel harmony processes are attested in Dagaare. These include tongue-root harmony and rounding harmony, with the latter interacting with both backness and height harmonies. These patterns are briefly discussed below.

3.1 Tongue root harmony: alternations in number suffixes

In Dagaare, a simple noun may consist of a noun root and suffix or just the noun root. Suffixes that occur with nouns mark number – singular and plural and they vary based on the noun class. A noun that is made up of only the root has all vowels agreeing in the same tongue root feature. If there is a number suffix, it alternates

4 It is worth noting that forms such as *bɪ́é* and *píɛ́* are often analyzed as bare roots without suffixes, with the final vowels /e/ and /ɛ/ considered epenthetic, inserted solely to satisfy prosodic minimality requirements. That is, the additional vowel is required to make up a bimoraic foot (Anttila and Bodomo 2009: 59–60). Under this view, these vowels function as phonological fillers with no morphological content or affiliation. However, this analysis is challenged when we examine noun–adjective compounds, such as: *bì-fáá* 'bad child' and *pɪ̀-ŋmáá* 'short roof'. In these examples, the final vowel is dropped in the presence of an adjective modifier, rendering forms like *bìè-fáá* or *pɪ̀ɛ̀-ŋmáá* ungrammatical. This pattern raises questions about the epenthetic interpretation and suggests that the final vowels may in fact function as suffixes, rather than meaningless insertions. Given this ambiguity, further phonological and morphological analysis may be necessary to definitively determine the status of these final vowels within Dagaare noun structures.

harmonically with the root. In the examples that follow, the suffix vowel is ATR when the preceding root is ATR (3a) and when the preceding root is RTR, the suffix is RTR as in (3b).

(3) Nominal suffixes

a. -V suffixes

ATR		RTR	
lı́ŋ-é	'lid-SG'	tı́ŋ-έ	'ground-SG'
tíg-é	'feast-PL'	sı́ŋ-ὲ	'waterpot-SG
bùg-ó	'silo-SG'	bʊ̀ŋ-ɔ́	'donkey-SG'
kóg-ó	'chair-SG'	pɔ́g-ɔ́	'woman-SG'

b. -CV suffixes

ATR		RTR	
jí-rì	'house-SG'	pı́ı́-rı̀	'sheep-PL'
bíí-rí	'child-PL'	sı́ı́-rı́	'waist-PL'
lín-ní	'lid-PL'	sı́n-nı̀	'waterpot'
tíg-rí	'feast-SG'	sı́g-rı́	'hut-SG'

The domain of tongue root harmony includes verbal morphology involving aspectual suffixes. This is given in the following section.

3.2 Tongue root harmony in verbs: alternation in suffixes

Verbal inflectional morphology in Dagaare simply consists of the verb root and an aspectual suffix. Selection of the verbform is determined by the aspectual context, which has been described in terms of perfective versus imperfective contrast (Anttila and Bodomo 1996; Saanchi 2003). The specific form of the aspectual suffix is controlled by phonological factors such vowel harmony, nasalization and tonal alternation. The verb form that has been described as perfective may consist of the verb root and the final high front vowel which is either [i] or [ɪ] depending on the tongue root feature of the vowels in the verb root. This final vowel may be optional. The verb may also be composed of just the root with no overt suffix. The final vowel never changes regardless of whether it marks aspect, infinitive or imperative, therefore suggesting that what has been described as perfective previously is in fact default morphology.

For the imperfective, there are always two overt morphemes, namely, the verb root followed by either a -V or a -CV suffix and this suffix always marks imperfective and nothing else. The quality of the suffix vowel is always determined by the vowel of the root. The following examples show Dagaare verbs and their morphological shapes.

(4) Aspectual suffixes

a. -V suffixes

i. Default forms

ATR		RTR	
búlí	'germinate'	bʊ̀rı̀	'sow'
dúórì	'urinate'	kʊ̀ɔ̀rı̀	'sell'
hə́ə̀rì	'yawn'	háárı̀	'shoo'

ii. Imperfective form

ATR		RTR	
búl-ò	'germinate-IPFV'	bʊ̀r-ɔ́	'sow- IPFV'
dúór-!ó	'urinate- IPFV'	kʊ̀ɔ̀r-ɔ́	'sell- IPFV'
hə́ər-!ə́	'yawn- IPFV'	háár-!á	'shoo- IPFV'

b. -CV suffixes

i. Default forms

ATR		RTR	
kpìnnì	'extinguish'	pɛ̀nnì	'rest'
ɲùnnì	'smell'	zànnì	'learn'
lònnì	'hurry'	mánnì	'measure/explain'

ii. Imperfective forms

ATR		RTR	
dì-ré	'eat- IPFV'	dí-!rɛ́	'take- IPFV'
túú-rò	'dig- IPFV'	tʊ́ʊ́-rɔ̀	'insult- IPFV'
dúó-rò	'climb- IPFV'	dʊ́ɔ́-rɔ̀	'weed- IPFV'
síg-rè	'descend- IPFV'	sɛ́g-rɛ̀	'write- IPFV'

Observe that aside the agreement in tongue root quality between the suffix and verbal root, there is also agreement in rounding/backness and low feature between the imperfective suffix and the verbal root. Next, is a description of rounding harmony.

3.3 Rounding and other harmonies[5]

Rounding harmony, or labial harmony, is a linguistic phenomenon where vowels within a domain share lip rounding – either all are rounded or all are unrounded (Kaun 2004, 1995). Common globally, it is most prominent in languages of Central Asia, Siberia, and Mongolia, but also appears in North American languages and several African Niger-Congo languages, including Dagaare, Chumburung, Igbo, Akan, and Nawuri (Bodomo 1997; Casali 1995a; Dolphyne 1988; Krämer 2003)

Languages with rounding harmony often also harmonize other features, like backness, as seen in Ewe (Odden 1991), or height, as in Nawuri (Casali 1995b). Some languages, such as Fante, combine rounding with tongue root harmony, making rounding harmony frequently dependent on other vowel features like height or backness (Kaun and McCollum 2024; Steriade 1981).

In Dagaare, rounding harmony mainly appears in imperfective verb forms (Bodomo 1997) and agentive reduplication, with fewer occurrences in nouns (Angsongna 2023). Consistent with Rose and Walker (2011), the imperfective in Dagaare harmonizes rounding with both backness and tongue root features: vowels are either all back and round or front and unround within a domain, and tongue root harmony is maintained across vowels, adding to the complexity of the harmony system.

3.3.1 Rounding harmony in verbs: alternations in aspectual suffixes

As stated earlier, the default verb (so-called perfective) form always has a final high vowel and this regardless of what shape the root vowel assumes. This suggests that default suffix or final high front vowel is never sensitive to roundness and backness even if it harmonizes with the root in terms of tongue root feature. This shows that the so-called perfective suffix is default morphology as the vowel remains unchanged in any conceivable environment.

The imperfective suffix always has a nonhigh vowel which always assimilates the preceding root vowel in terms of rounding, backness and tongue root, and height feature. Effectively, when the verb root contains a round back vowel, the vowel of the imperfective suffix is also round back as in (5a) and when the vowel in the verb root is

5 Aside Central Dagaare, rounding is also reported in reduplication in Birifo (Dundaa 2013), a related variety of Dagaare. It is reportedly absent in Dagara (Kuubezelle 2013) and Waale (Abdul Moomin 2015). An anonymous reviewer, while acknowledging that rounding harmony is not a prominent feature of Dagara-Lobr (a dialect of Dagara), notes that it does appear in certain contexts – specifically in an allomorph of the number/noun class suffix -rV found in some nouns. For example: dʊɔwr (sg) - dɔrʊ (pl) 'yellow', nuu (sg) – nuru (pl) 'hand'.

unround and front, the vowel of the suffix assumes identical features, as shown in (5b). In (5c), while agreeing in tongue root feature, a root with low vowel always combines with a suffix with a low vowel.

(5) Imperfective suffixes

a. [round mid]

wùl-ó	'show- IPFV'	sɔ́r-ɔ̀	'count/read-IPFV'
túú-rò	'dig- IPFV'	tʊ́ʊ́-rɔ̀	'insult-IPFV'
dúó-rò	'climb- IPFV'	dʊ́ɔ́-rɔ̀	'weed-IPFV'

b. [unround mid]

jíél-!é	'sing- IPFV'	jíɛ́l-!ɛ́	'winnow-IPFV'
dì-ré	'ear- IPFV'	dí-!rɛ́	'take-IPFV'
lé-rè	'tie- IPFV'	síɛ́-rɛ̀	'roast- IPFV'

c. [unround low]

hə́ə́r-!ə́	'yawn- IPFV'	háár-!á	'shoo-IPFV'
sə́ə́r-!ə́	'gather- IPFV' (in bulk)'	záá-rà	'throw- IPFV'
ɡə́ə́r-!ə́	'belch/burp- IPFV'	dáá-!rá	'push- IPFV'

Based on the data, both tongue root and rounding harmony propagate from the root outward to the suffix or final vowel. Thus, Dagaare harmony can be considered as root-controlled (Clements 1985; Kirchner 1993). With the preceding exposition to the main types of vowel harmony in Dagaare, let us now turn our attention to a proposed analysis.

4 Computation and phonology

Various theoretical frameworks have been developed to explain phonological phenomena like vowel harmony. Key among these are Optimality Theory (OT) (Mccarthy 2011; McCarthy and Prince 1993, 1999; Prince and Smolensky 2004), Emergent Phonology (Archangeli and Pulleyblank 2017, 2022), and Autosegmental Phonology (Goldsmith 1976), Generative Phonology (Clements 1980; Kenstowicz and Kisseberth 2023; Roca 2003) and many others, with OT emerging as one of the most widely used approaches. In this paper, I apply a formal analysis within the computational theory framework (Aksënova et al. 2024; Heinz 2011, 2018; Heinz and Lai 2013). Since the early twentieth century, computational theory has driven some of the most influential scientific advances, primarily through its ability to formally define problems, processes, and computational solutions, providing a rigorous basis for examining phonological patterns and processes.

In linguistics – the scientific study of language – exploring cognitive processes is central to understanding the human capacity for language. In phonology, this exploration involves examining the nature of phonological representations, constraints, and processes in phonological grammar. As Aksenova et al. (2024) highlight, the key questions in phonology include: (i) What constitutes knowledge of phonological well-formedness? (ii) What representations are used to determine this knowledge? (iii) How is this knowledge acquired? (iv) What types of functions underlie phonotactics and phonological processes? and (v) What data structures are involved in these processes? Each of these questions represents a unique computational problem, providing a foundation for further investigation into the workings of phonological cognition.

This paper seeks to address key computational questions in phonology by defining necessary conditions for vowel harmony in Dagaare, specifically through the framework of Tier-based Strictly Local (TSL) grammars. Generally, a process or problem is considered computable if an algorithm can consistently produce correct outputs from any given input. In linguistic theory, grammars are viewed as algorithmic processes with distinct roles.

Two types of grammars are relevant here: phonotactic grammar, which evaluates the well-formedness of surface structures (e.g., sequences of speech sounds), and phonological grammar, which maps underlying forms to surface structures. While phonotactic grammar assesses surface well-formedness directly, phonological grammar derives new structures from inputs. This study adopts two computational perspectives: one examines the constraints governing phonotactic patterns in Dagaare vowel harmony (involving tongue root and rounding/backness and height harmonies), and the other investigates the constraints governing harmony processes. Using TSL grammar, which selectively projects relevant elements to a tier while disregarding irrelevant ones, this approach enables evaluation of surface forms to identify constraint violations. This method offers a clear computational model for analyzing vowel harmony in Dagaare.

4.1 Formal languages

First, I give a brief review of Strictly Local grammar in 4.1.1 and how it fails to capture long distance patterns. Then I turn to Tier-Based Strictly Local in Section 4.1.2.

4.1.1 Strictly local grammars

Strictly Local (SL) grammars are a class of formal languages that serve as the foundation for modeling strictly local dependencies in phonological patterns (Chandlee 2014). Within computational phonology, SL grammar defines phonological well-formedness based on local constraints over contiguous substrings of length *k*, known as *k-grams*. A string is considered well-formed with respect to a constraint set *C* if it does not contain any forbidden *k*-grams (Graf 2017). In essence, SL grammar characterizes acceptable strings by excluding certain prohibited substrings, making locality the central organizing principle. Formally, a SL grammar consists of the set of *k-grams*, and G_{SL} that must not be contained within a well-formed string of the language (Aksënova et al. 2024).

For example, in Samala (a Chumashan language spoken in California's Santa Ynez Valley), there is a local dissimilation process whereby /sn/, /sl/, and /st/ are realized as [ʃn], [ʃl], and [ʃt], respectively (McMullin 2016). This can be interpreted as a phonotactic restriction that prohibits [s] from being immediately followed by [n], [l], or [t]. Such restrictions are straightforwardly captured using Strictly Local (SL) grammars, which encode well-formedness by banning specific bigrams, that is, sequences of two adjacent segments – in this case, *sn*, *sl*, and *st*.

Another illustration comes from Russian, where obstruent clusters must agree in voicing. Based on the analysis in Aksënova et al. (2024), this voicing agreement constraint is local and can be captured using SL grammars by banning sequences like [+voice] [-voice] or [-voice] [+voice] in obstruent clusters, as illustrated below.

(6) Voicing in Russian obstruents
   a. [s]-tajat^j ‘PERF-melt’ *[z]-tajat^j
   b. [z]-bit^j ‘PERF-beat’ *[s]-bit^j

The telic prefix *s-* is realized as [s] when followed by a voiceless consonant (see 6a), and as [z] when followed by a voiced consonant (see 6b). The relevant alphabet can be defined as ∑ = {s, z, t, b, j, a, i}. The phonotactic grammar in this case prohibits sequences of obstruents with mixed voicing, i.e., adjacent obstruents that disagree in their voicing specifications. This can be formalized as a Strictly Local grammar: G_{SL} = ⟨*sz, *zs, *sb, *bs, *zt, *tz, *tb, *bt⟩. This grammar correctly rules out ill-formed sequences where the voicing of the prefix *s-* does not match that of the following obstruent. For instance, forms where a voiceless [s] precedes a voiced consonant (e.g., *sb, *sz), or where a voiced [z] precedes a voiceless consonant (e.g., *zt, *zs), are predicted to be ungrammatical under this system.

In the context of vowel harmony, SL grammars are used to model patterns where harmonic features such as [±ATR], [±back], or [±round] must match across adjacent or nearby vowels within a word. The core principle of SL grammar is the specification of permitted or prohibited local sequences (typically pairs of segments, or *bigrams*) in a phonological string. A string is considered well-formed if it does not contain any of the prohibited sequences. SL grammar is especially well-suited for modeling vowel harmony within single morphemes or short stems, where harmony is strictly local. For example, in Dagaare, vowels within roots and suffixes typically agree in [±ATR] (Advanced Tongue Root), as illustrated in (7) below.

(7) Dagaare tongue root harmony
ATR RTR
a. túò 'baobab.SG' bʊ́ɔ́ 'goat.SG'
b. zíé 'place.SG' nɪ́ɛ̂ 'person.SG'
c. bùg-ó 'silo-SG' bʊ̀ŋ-ɔ́ 'donkey-SG'

The vowels involved in the harmony pattern are either directly adjacent or separated by a single intervening consonant. Crucially, they must agree in their tongue root feature – either [ATR] or [RTR]. The relevant vowel alphabet is ∑ = {i, e, u, o, ɪ, ɛ, ʊ, ɔ}, and the SL grammar rules out combinations where [ATR] and [RTR] vowels co-occur locally. Specifically, the following bigrams are prohibited: G_{SL} = ⟨*iɛ, *ɛi, *ɪe, *eɪ, *uɔ, *ɔu, *ʊo, *oʊ⟩. These sequences represent ill-formed vowel pairings that violate tongue root harmony and are correctly excluded by the SL bigram constraints.

Now let us look at another harmonic process. In Lokaa (as discussed in Akinlabi 2009; Aksënova et al. 2024), non-high vowels must harmonize in ATR value with the preceding non-high vowel, while high vowels and consonants are transparent to harmony. This pattern results in both well-formed and ill-formed surface forms, with agreeing segments that may not be adjacent to each other. The following examples illustrate this harmonic agreement, highlighting instances where agreement is achieved across intervening segments.

(8) Lokaa harmony
a. èsìsòn 'smoke' *èsìsɔ̀n
b. ɛ̀sísɔ̀n 'housefly' *ɛ̀sísòn
c. lèjìmə̀ 'matriclan' *lèjìmà
d. ɛ́kílìkà 'kind of plant' *ékílìkà

The strings in (8a) and (8b) need 5-grams to capture the pattern as there are three intervening elements between the two non-high vowels, [e] and [o]. In (8d), there are five segments intervening between [ɛ] and [a] and because there is no upper bound on the amount of material which separate the two non-high vowels agreeing in terms of [ATR] feature, no SL grammar, which depends on a fixed locality window, can adequately account for this pattern. This limitation highlights the need for Tier-Based Strictly Local (TSL) grammar, which can handle long-distance dependencies by projecting only harmony-relevant segments onto a separate tier. In short, while SL grammars are effective for modeling strictly local patterns, they fail to capture non-local vowel harmony, especially when intervening consonants or syllables must be ignored. This necessitates the use of more expressive models, such as TSL, to accurately represent these long-distance phonological processes.

4.1.2 Tier-based strictly local grammars

Tier-based Strictly Local (TSL) grammars (Aksënova et al. 2024; Heinz 2011, 2018; Heinz et al. 2011) enable the modeling of non-local dependencies by projecting specific input elements onto a tier, creating locality among otherwise distant segments and ruling out forbidden strings. TSL grammar consists of a tier alphabet, *T*, and a set of k-grams, G_{TSL}, that must not appear in a well-formed tier representation of a string. This approach allows TSL grammar to block forbidden substrings over the tier, even when the segments are not contiguous in the linear string. For example, the harmony pattern in Lokaa in (8) shows non-high vowels harmonizing across intervening segments. Although this pattern is impossible to capture with purely SL grammar, TSL enables projection of these non-adjacent segments onto a tier, where ill-formed combinations can be effectively ruled out. Thus, Table 1 below shows the active constraints on tier-adjacent non-high vowels. The ATR pattern in the surface forms is obtained by blocking the combinations of non-high vowels that do not agree in [tense] specifications.

There are many languages in which segments agree in more than one single feature. For instance, in Turkish, vowels agree in backness and rounding (see, Kaun 1995, 2004); in Bukusu, one agreement pattern involves vowels and another involves liquids (Odden 1994); in Dagaare verbs, the imperfective suffix agrees in ATR and rounding

Table 1: TSL grammar for Lokaa harmony.

	Tier of non-high vowels T = {ɛ, e, o, ə, ɔ, a}
1.	*[α tense] [- α tense]
	H_{ATR} = {*ɛe, *eɛ, *oɛ, *oɛ, *ɛə, *əɛ, *ɔe, *eɔ, *ɔo, *oɔ, **aə, *ao, *ae, *ea}

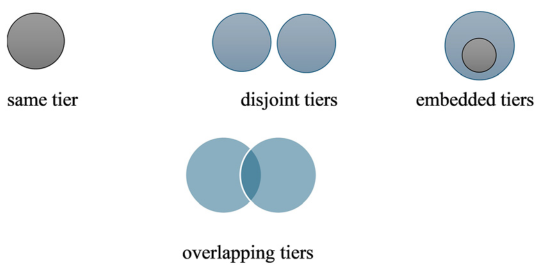

Figure 2: Possible tier relations.

with the verb root (Angsongna 2023). Such languages with more than one harmony pattern may require more than one tier and these tiers may be related in some ways. Figure 2 illustrates the four logically possible relations between tier 1 and tier 2 (Aksënova et al. 2024): (i) T1 and T2 can be identical as in double harmony involving only vowels; the set of blockers and undergoers is the same for both harmony processes (same tier) (ii) T1 might share no common elements with T2 – disjoint tiers. This relation involves independent vowel harmony and consonant harmony processes (iii) embedded relation in which T1 is a subset of T2 or vice versa. This is the case of consonant harmony such as sibilant harmony (iv) partially overlapping relation in which T1 and T2 share some elements, but neither is a subset of the other.

Cases where T1 is the only tier that partially overlaps T2 appear to be typologically unattested. Meanwhile, out of the four logical possibilities, the partially overlapping case is reported to be the most common.

TSL approach is used to capture different harmonic patterns. A study of double vowel harmony patterns shows they require a single tier, although more than one feature is involved. For sibilant harmonies however, two tiers are needed. Languages with both vowel and consonant harmonies also require two tiers though the two tiers do not share any common element. That is, such cases involve disjoint tiers.

With double vowel harmonies, the initial assumption might be that each harmonic feature requires a separate tier. However, data reveals that both harmonies can coexist on the same tier. This does not imply that undergoers and blockers are identical for both harmonies; each segment participating in one harmony remains relevant to the other. Typically, one harmony affects all vowels, while the other spreads only among tier-adjacent vowels. In some instances, both features spread simultaneously, while in others, all vowels act as undergoers for both harmonies, with a subset further propagating both features. Essentially, TSL grammar formalizes a particular constraint within a language phonological structure. In the following sections, I analyze the multiple vowel harmonies observed in Dagaare using TSL grammar as the framework.

4.1.3 Tongue root harmony in TSL

Tongue root harmony in Dagaare as described in previous sections affects all vowels. As a result, TSL grammar then allows the projection of dependent items on the tier and bans all ill-formed combinations over the tier. There are two harmonizing features in Dagaare that are either plus or minus and therefore all vowels within a simple word have a choice of the following two possibilities: [ATR] – [ATR] and [RTR] – [RTR]. For each of these combinations, there is an example below which shows that suffixes must agree with the root in these parameters.

(9)

	ATR			RTR		
	tíg-é	'feast-PL'	*tíg-ɛ́	tíŋ-ɛ́	'ground- SG'	*tíŋ-é
	dúŋ-ó	'animal- SG'	*dúŋ-ɔ́	dʊ́ŋ-ɔ́	'mosquito- SG'	*dʊ́ŋ-ó
	kóg-ó	'chair- SG'	*kóg-ɔ́	pɔ́g-ɔ́	'woman- SG'	*pɔ́g-ó
	nəɲìgé	'thief-sg'	*nəɲìgɛ	zɔ̀mìŋɛ́	'friend-sg'	*zɔ̀mìŋé

The data above and harmony requirements of the language give rise to the following well-formedness constraints as adopted from Pulleyblank (2002).

(10) i. *ATR-Co-RTR: Ignoring consonants, an ATR segment may not be followed by an RTR segment
ii. *RTR-Co-ATR: Ignoring consonants, an ATR segment may not be followed by an ATR segment

The constraints prohibit vowel sequences where the tongue root features differ. Table 2 shows the active constraints on the tier adjacent to the tier alphabet. The constraints illustrate the illicit featural combination. The tier alphabet includes all vowels in the language. In addition, there are corresponding ill-formed segmental bigrams which make up the forbidden substrings in the TSL grammar H_{ATR}. Any bigram in which the vowels disagree in tongue root feature must be banned.

Figure 3 below illustrates the analysis above. The subfigure in (a) shows the well-formed word *tígé* 'feasts' while the one in (b) shows the illicit word.

The ATR vowels [i, e] are projected on the tier and their combinations *i.e.* is not among the sequences that are required to be ruled out and therefore the word *tígé* is well-formed in (a). The other subfigure in (b) shows ill-formedness with the word **tígɛ́*, where the two adjacent vowels *í* and *ɛ* disagree in their tongue root

Table 2: Dagaare Tongue root harmony.

	Vowel tier T= {i, e, u, o, ə, ɪ, ɛ, ʊ, ɔ, a}
1.	*ATR-Co-RTR *RTR-Co-ATR
	H_{ATR} = {*iɪ, *ɪi, *iɛ, *ɛi, *iʊ, *ʊi, *iɔ, *ɔi, *ia, *ai, *eɪ, *ɪe, *eɛ, *ɛe, *eʊ, *ʊe, *eɔ, *ɔe, *ea, *ae, *uɪ, *ɪu, *uɛ, *ɛu, *uʊ, *ʊu, *uɔ, *ɔu, *ua, *au, *oɪ, *ɪo, *oɛ, *ɛo, *oʊ, *ʊo, *oɔ, *ɔo, *oa, *ao, *əɪ, *ɪə, *əɛ, *ɛə, *əʊ, *ʊə, *əɔ, *ɔə, *əa, *aə

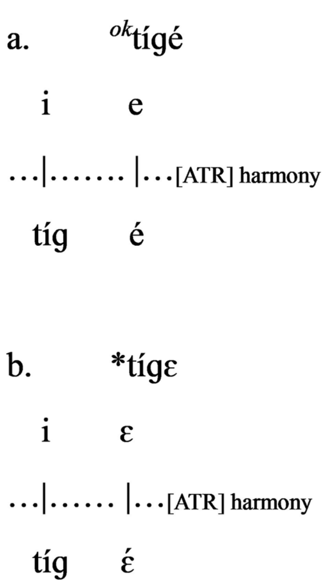

Figure 3: Tongue root harmony in Dagaare.

specifications. Accordingly, when *i* and *ɛ* are projected on a tier, we see that the bigram **iɛ* is prohibited by the grammar and so the word **tigɛ* must be ruled out.

It is important to note that the notion of directionality – such as root/stem control – does not directly factor into the implementation of TSL grammars. Rather, TSL focuses on the sequencing of relevant segments (e.g., vowels), without reference to morphological structure. In this framework, directionality is not explicitly encoded but instead emerges from the nature of the prohibited featural combinations specified in the grammar.

Overall, TSL grammars are particularly well-suited for capturing the phonotactic patterns of vowel sequences in simplex words, especially in languages like Dagaare, where vowels within a word must harmonize in a specific harmonic feature (e.g., tongue root) specification. However, this effectiveness diminishes when applied to morphologically complex structures such as compounds, which present significant challenges for the TSL framework due to its lack of morphological sensitivity. I return to this issue in Section 4.2.

The following section, however, shifts focus to how TSL can be extended to model multiple types of vowel harmony – specifically those involving tongue root, rounding, backness, and height features.

4.1.4 Multiple vowel harmonies in TSL (ATR, rounding, backness, and height)

As previously stated, with TSL grammar, a harmonic process selects a set of elements (undergoers) and establishes an agreement relation among them with respect to a particular feature. These elements are projected on the tier alphabet T (which contains all the vowels of the language) in order to achieve locality. In the previous section, using TSL grammar, I give an account of tongue root harmony, which is the main harmonic pattern exhibited in nouns. In this section, I present a formalization of the harmony patterns involving the imperfective form of the verb. The imperfective form in Dagaare, unlike nouns and the default verb forms, exhibits cases where vowels agree in more than one harmonic feature – tongue root, rounding, backness and height. This multiple feature agreement is commonly observed in many languages. For instance, in Turkish, vowels agree in rounding and backness; in Imdlawn Tashlhyt sibiliants agree in voicing and anteriority and in Bukusu one agreement pattern involves vowels and another involves liquids. These are cases often described as double harmonies. These patterns raise questions as to whether more than one tier is required to account for double harmonies or whether there is a restriction on the tier alphabet for patterns that exhibit double harmonies.

Considering a single tier is required to account for the single harmonic pattern ATR, it is likely to easily assume that each feature of a double or multiple harmony needs its own tier. That is not the case though, because a typological study (Aksënova et al. 2024; Burness et al. 2021) reports that double vowel harmony requires just a single tier despite the presence of more than one feature. All harmonies fit on the same tier. It should be noted that the undergoers and blockers are not the same for both or all harmonies but the elements taking part in one harmony are relevant for the other because sometimes the double or multiple harmonic features are transmitted simultaneously.

The imperfective form in Dagaare involves multiple harmonies (tongue root, rounding, backness and height). All harmonic features in these patterns are simultaneously transmitted. As shown in the following data (11), within an imperfective verb, all vowels agree in the above features – the imperfective suffix always agrees in terms of the above features with the root.

(11) a. Imperfective suffixes
[front unrounded mid]
jíél-!é ‘sing-IPFV’ jíɛ́l-!ɛ́ ‘winow- IPFV’
dì-ré ‘eat- IPFV’ dí-!rɛ́ ‘take- IPFV’
b. [back rounded mid]
wùl-ó ‘show/teach- IPFV’ sɔ́r-ɔ̀ ‘count/read- IPFV’
túú-rò ‘dig- IPFV’ tʊ́ʊ́-rɔ̀ ‘insult- IPFV’
c. [urounded low]
hə́ə́r-!ə́ ‘yawn- IPFV’ háár-à ‘shoo- IPFV’
sə́ə́r-ə̀ ‘gather- IPFV (in bulk)’ záá-rà ‘throw- IPFV’

These examples in (11) and the requirements for vowel harmony in Dagaare leads to the constraint, AGREE (F). This constraint is further expanded into the following three-way set.

(12) a. AGREE (ATR): Adjacent vowels in a domain agree in the same value for the feature [ATR]
b. AGREE (RND): Adjacent vowels in a domain agree in the same value for the feature [round]
c. AGREE (LOW): Adjacent vowels in a domain agree in the same value for the feature [low]

Similar to the tongue root harmony discussed above, the four harmonizing features for the imperfectives can be plus or minus. Therefore, all vowels within an imperfective verb have a choice among the following possibilities: [±ATR], [±round], [±back], [±low]. Table 3 below presents the TSL grammar for Dagaare imperfective verbs in various harmonic components. The tier alphabet *T* incudes all the vowels in Dagaare. H_{ATR} rules out sequence of vowels that disagree in tongue root feature; H_{round} prohibits co-occurrence of vowels that do not agree in rounding feature; H_{back} bans vowels in a sequence that do not share same feature for backness and finally H_{low} rules out vowels in a domain that do not share same feature low. In effect, only one TSL grammar is necessary – its tier alphabet *T*, and $G_{TSL} = H_{ATR} \cup H_{\text{round}} \cup H_{low}$. That is, this grammar is basically a three-way conjoined constraint of AGREE (F) banning all V-V sequences that disagree in one or more of these three features.

Only one TSL grammar is necessary to capture these patterns. This grammar operates within the tier alphabet *T* and the grammar G_{TSL} is a combination of H_{ATR}, H_{round}, and H_{low} which contain the sets of forbidden bigrams. Figure 4 presents a graphic representation of the analysis above. In each subfigure (a–c), the well-formed word appears on the left, with its corresponding illicit form on the right.

The subfigures in (a) represent ATR harmony (though it exhibits backness/rounding as well). The left subfigure represents the well-formed word, *wùló*. The two vowels are projected on a tier, and nothing is blocked

Table 3: TSL grammar for rounding/backness, tongue root and height harmony.

	Vowel tier T={ i, e, u, o, ə, ɪ, ɛ, ʊ, ɔ, a}
1.	*[αATR] [βATR] H_{ATR} = {*iɪ, *ɪi, *iɛ, *ɛi, *iʊ, *ʊi, *iɔ, *ɔi, *ia, *ai, *eɪ, *ɪe, *eɛ, *ɛe, *eʊ, *ʊe, *eɔ, *ɔe, *ea, *ae, *uɪ, *ɪu, *uɛ, *ɛu, *uʊ, *ʊu, *uɔ, *ɔu, *ua, *au, *oɪ, *ɪo, *oɛ, *ɛo, *oʊ, *ʊo, *oɔ, *ɔo, *oa, *ao, *əɪ, *ɪə, *əɛ, *ɛə, *əʊ, *ʊə, *əɔ, *ɔə, *əa, *aə
2.	*[α round] [β round] H_{round} = {*iʊ, *ʊi, *iɔ, *ɔi, *eʊ, *ʊe, *eɔ, *ɔe, *uɪ, *ɪu, *uɛ, *ɛu, *ua, *au, *oɪ, *ɪo, *oɛ, *ɛo, *oʊ, *ʊo, *oɔ, *ɔo, *oa, *ao, *əʊ, *ʊə, *əɔ, *ɔə,
3.	*[α low] [β low] H_{low} = {*ia, *ai, *ea, *ae, *ua, *au, *oa, *ao, *əɪ, *ɪə, *əɛ, *ɛə, *əʊ, *ʊə, *əɔ, *ɔə,

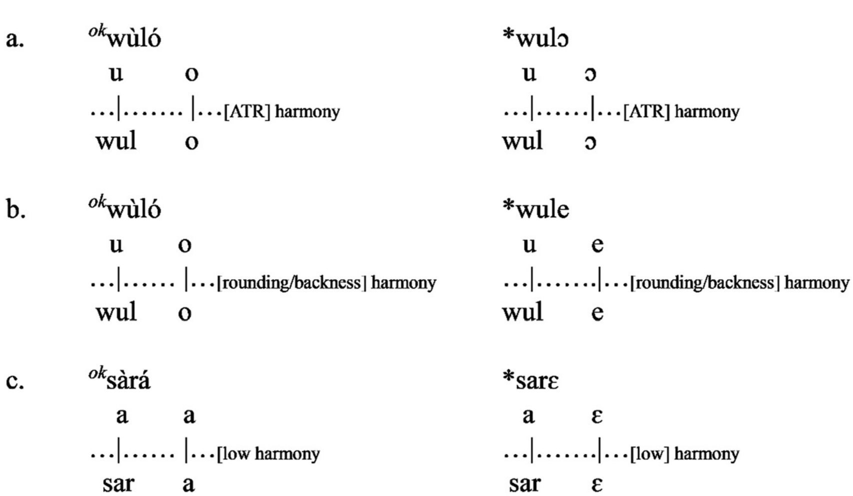

Figure 4: ATR, rounding/backness, and height harmony in Dagaare.

because the combination *uo* is acceptable as they both agree in ATR features. The right figure in (a) represents a violation of tongue root harmony where **uɔ* is ruled out by the H_{ATR} part of the grammar. The left subfigure of (b) is a graphic representation of rounding/back harmony where nothing is blocked while the subfigure on the right illustrates disagreement in round/back harmony wherefor the bigram **ue* is avoided by the H_{round} of the grammar. Finally, the subfigures in (c) demonstrate height harmony in which *aa* is not blocked since there is agreement is the feature [low]. The figure on the right however is illicit because the combination of the low and mid vowels **aɛ* is ruled out by the H_{low} component of the grammar. In summary, the harmonic patterns in the imperfective involve cases where the undergoers for one harmonic spreading is the same for the others, hence only a single TSL grammar is required to capture these patterns.

Overall, Dagaare exhibits multiple harmony processes that are effectively captured within a single Tier-Based Strictly Local (TSL) grammar framework. This unified TSL grammar is essential because each vowel participating in one harmony is relevant to the others; the harmonic processes are interdependent rather than isolated. TSL grammar serves as a formal mechanism for encoding specific phonotactic constraints that govern permissible and impermissible surface sequences within a language. It accomplishes this through either individual constraints or a set of constraints acting in conjunction, but it does not account for underlying phonological alternations that arise when roots, suffixes, and clitics are combined. Thus, TSL grammar is known for its efficiency in modeling surface phonotactics directly, without relying on underlying forms, a feature that distinguishes it from many generative phonological models such as Optimality Theory (OT). It focuses exclusively on surface-level segmental sequences, allowing it to characterize observable harmonic patterns with precision.

Crucially, the TSL account of Dagaare multiple harmonies – specifically tongue root, rounding, backness, and height – is comparable to the double vowel harmony systems observed in languages such as Kirgiz, Yakut, and Buryat, as discussed by Aksënova et al. (2024). In all these cases, the harmonies target the same set of segments (i.e., vowels), allowing a single TSL grammar to capture all the harmony patterns simultaneously. This makes Dagaare typologically similar to these languages in terms of its representational simplicity under the TSL framework.

However, Dagaare differs from languages like Imdlawn Tashlhiyt (Berber), which exhibits sibilant harmony involving both anteriority and voicing, and Kikongo (Bantu), which displays vowel harmony (in height) alongside nasal harmony. In these cases, the harmonies affect distinct sets of segments (e.g., sibilants and nasals) and thus cannot be captured within a single tier. As a result, each harmony pattern in these languages requires a separate TSL grammar, with distinct tiers for each type of harmony. This contrasts with Dagaare, where a single tier suffices for modeling multiple vowel harmonies, highlighting the representational efficiency of its harmonic system under the TSL framework.

4.2 Lack of morphological boundary sensitivity

As mentioned before, TSL grammars are especially effective in accounting for the phonotactics of vowel sequences in simplex words, where vowels within a simple word must agree in a harmonic feature (e.g., tongue root, rounding) specifications.

However, this constraint appears to be relaxed in certain compound structures. Some vowel combinations considered illicit in simplex words do occur in compounds, which TSL grammars fail to account for due to their lack of morphological awareness. While TSL can model long-distance dependencies by projecting only harmony-relevant segments (e.g., vowels), it struggles when intervening transparent segments – those that do not participate in harmony but do not block it either – cannot be consistently projected (Chandlee 2014).

Moreover, vowel harmony in many languages is sensitive to morphological boundaries. In Turkish, for example, harmony typically applies within morphemes but not across compound or derivational boundaries (Gouskova 2018; Kabak and Vogel 2001). Similarly, in Dagaare, vowel harmony often does not extend across compound words. As (Chandlee 2014; Jurgec 2011) argue, accurate modeling of such phenomena requires domain-sensitive tier projection – a mechanism that standard TSL grammars lack.

As a result, TSL grammars are ill-equipped to handle compound-induced exceptions to vowel harmony in Dagaare. Their inability to recognize morphological boundaries limits their descriptive and predictive power. Consider the following Dagaare compounds.

(13) a. Noun-noun compounds

noun	gloss	noun	gloss	compound	gloss
bíé	‘child.SG’	dɔ́ɔ́	‘man.SG’	bì-dɔ́ɔ́	‘boy.SG’
bíé	‘child. SG’	pɔ́g-ɔ́	woman- SG’	bì-pɔ́g-ɔ́	‘girl- SG’
zû	‘head. SG’	kɔ́m-ɔ́	hair-PL’	zu-!kɔ́m-ɔ́	‘hair (on head)

b. Noun-adjective compounds

noun	gloss	adjective	gloss	compound	gloss
bíé	‘child.SG’	fáá	‘bad’	bì-fáá	‘bad child.SG’
kɔ̀lʊ́ŋ	‘well.SG’	zùlúŋ	‘deep’	kɔ̀lʊ̀ŋ-zùlúŋ	‘deep well.SG’
bíé	‘child. SG’	zɪ̀ɛ́	‘red’	bì-zɪ̀ɛ́	‘fair child. SG’

The examples above clearly demonstrate that all the compounds presented violate the tier-based restrictions imposed by TSL grammars. Sequences such as **iɔ*, **uɔ*, **ia*, **ɔʊu*, and **iɪɛ* are prohibited under the TSL framework, yet they are attested and well-formed in morphologically complex constructions like compounds. This highlights a core limitation of TSL: while it effectively captures a range of harmonic patterns in simplex words and accounts for many long-distance phonological dependencies (Chandlee 2014; Heinz 2018), it fails to accommodate the complexities introduced by morphological structure. Specifically, the presence of word or morpheme boundaries, such as the right edge of the first root and the left edge of the second in compounds, introduces opacity effects that disrupt the predictions made by TSL-based analyses.

These findings suggest that while TSL provides a robust model for local and even some non-local phonological processes, it lacks the structural sensitivity needed to handle cross-morphemic or morphologically complex phenomena. To address this limitation, further refinement of the framework is necessary – potentially through the incorporation of domain-sensitive tier projection or other representational extensions that account for morphological structure.

5 Learnability

Computational approaches such as Tier-Based Strictly Local (TSL) grammars are instrumental in linking linguistic patterns to learnable algorithms. TSL grammar provides structured hypotheses for deriving grammars from empirical patterns, supporting the perspective of Heinz et al. (2022), who view phonological learning as a function mapping data to structured rules. As Lambert et al. (2021) note, this framework is efficiently learnable, framing learner success in terms of reliably extracting grammar from surface data.

This study contributes to understanding how vowel harmony, particularly in Dagaare, can be learned by both human language learners and computational systems by integrating empirical phonological description with a computationally grounded formal model. This combination of descriptive analysis and formal representation makes harmonic patterns more transparent, learnable, and generalizable.

One of the primary learnability gains comes from formulating vowel harmony in terms of explicit, observable surface constraints rather than abstract derivational rules. As discussed, TSL grammars define well-formedness in terms of permissible or prohibited local sequences (e.g., bigrams) on a relevant tier (Chandlee 2014; Heinz 2010). In Dagaare, for instance, the [±ATR] harmony constraint in simplex words can be captured by banning illicit combinations such as *iɛ, *uɔ, and *ʊo. This aligns with learnability research (Heinz 2010; Heinz and Idsardi 2013) showing that surface-true generalizations are more easily acquired from positive evidence, as they do not require hypothesizing abstract underlying forms. For learners, these patterns are directly detectable in input, avoiding the need for complex morphophonemic reasoning.

While Strictly Local (SL) grammars are limited to adjacent dependencies, many vowel harmony systems require capturing non-local dependencies. TSL addresses this by projecting only harmony-relevant segments (e.g., vowels) while ignoring intervening consonants or transparent vowels (Heinz et al. 2011). This reduces the learner's memory and processing load by filtering out irrelevant material. Jardine and McMullin's (2017) findings directly support this approach: they show that TSL grammars can be learned in polynomial time from positive evidence alone, and that learners can infer the relevant tier without prior specification while disregarding irrelevant segments. This efficiency parallels the Dagaare case, where projecting only vowels allows a single TSL grammar to capture multiple vowel harmonies ([±ATR], [±round], [±back]) without overgeneralizing to compounds or morphologically complex forms – patterns that TSL correctly excludes.

Importantly, Dagaare multiple harmonies – [±ATR], [±round], [±back] – can all be modeled within a single TSL grammar, unlike in languages such as Kikongo or Imdlawn Tashlhiyt, where multiple harmonies require separate TSL tiers (Aksënova et al. 2024; Jurgec 2011). From a learnability perspective, a single formal analysis reduces the hypothesis space, allowing learners to apply one mechanism across multiple features rather than mastering separate rules for each.

The study also shows that while TSL effectively models harmony in simplex words, it fails in compounds and morphologically complex forms, where illicit sequences are tolerated. This has direct learnability implications: recognizing such morphological boundaries prevents overgeneralization, a common error in both human and machine learning (Gafos and Dye 2011). A model sensitive to morphological domains thus enables learners to apply harmony rules only where appropriate.

Finally, while generative approaches such as Optimality Theory, Harmonic Grammar and others provide rich morphophonemic explanations, their learning models require inferring abstract, ranked constraints from data (Tesar and Smolensky 2000). In contrast, TSL:

- operates solely on attested surface strings,
- uses a finite-state representation provably learnable from positive data, and
- avoids the complexity of hidden structures, which experimental and computational studies (Jarosz 2013) have shown to impede acquisition.

For learners, this means the Dagaare harmony system, formalized through TSL, presents itself as a finite, surface-consistent set of constraints that is easier to acquire and generalize than abstract ranking-based systems.

6 Discussion and conclusion

This study applies the Tier-Based Strictly Local (TSL) grammar framework to model vowel harmony in Dagaare, examining multiple harmony types – tongue root, rounding, backness, and height. The analysis demonstrates that TSL grammars effectively capture the surface phonotactics of harmonic patterns in simplex words by encoding constraints as permissible or prohibited local sequences on a relevant tier. By projecting only vowels, the model successfully represents non-local dependencies while ignoring intervening consonants. This strength mirrors the findings of Jardine and McMullin (2017), who showed that TSL languages can be learned from positive evidence without prior tier specification, reinforcing the cognitive and computational plausibility of the approach.

A notable strength of the analysis is that all harmony types in Dagaare can be represented within a single TSL grammar. This single-tier approach reduces the learner's hypothesis space and simplifies acquisition, as one mechanism applies across multiple features. Furthermore, the study confirms that TSL captures surface-true generalizations, aligning with learnability research showing that such patterns are more readily acquired than those requiring abstract underlying rules.

However, the findings also reveal a key limitation: TSL grammars fail to account for harmonic exceptions in morphologically complex words, particularly compounds, where vowel sequences prohibited in simplex words are nonetheless attested. This shortcoming arises from the lack of morphological domain sensitivity of the framework. Morphological boundaries such as those between two roots in a compound can disrupt harmony, producing opaque surface forms that the model cannot predict without additional representational machinery, such as domain-sensitive tier projection (Jurgec 2011). Without such refinements, TSL risks overgeneralization when applied indiscriminately across word types.

In conclusion, this study shows that TSL grammar provides a robust, learnable, and efficient framework for modelling vowel harmony in Dagaare, particularly in morphologically simple forms. Its ability to capture multiple harmony types within a single representation reflects the typological profile of Dagaare and supports cognitive plausibility by minimizing processing and memory demands for learners. TSL thus offers a compelling alternative to traditional phonological theories for analyzing complex harmony systems, with clear applications in both linguistic theory and computational phonology. Nonetheless, the inability of standard TSL to incorporate morphological boundaries highlights a direction for future research. Extending TSL with domain-sensitive tiers or hybrid models that integrate finite-state surface constraints with morphological awareness could enhance its capacity to model harmony in morphologically complex contexts. Beyond Dagaare, the findings invite replication of this approach in related languages to further assess its theoretical and typological utility.

Abbreviations

1SG	= First person singular
1PL	= First person plural
2SG	= Second person singular
IPFV	= Imperfective
NMLZ	= Nominalizer
NEG	= Negation
SG	= Singular
PL	= Plural

Acknowledgements: I would like to express my gratitude to the native speakers of Dagaare in Sombo (Nadowli-Kaleo district) for generously sharing their time and language with me during my fieldwork in Ghana.
Conflict of interest: The author states no conflict of interest.
Funding information: The data for this study was collected through fieldwork funded by an Insight Grant (#435-2016-0369) from the Social Sciences and Humanities Research Council of Canada (SSHRC), awarded to my PhD supervisor, Douglas Pulleyblank.
Author contribution: The author confirms the sole responsibility for the conception of the study, presented results and manuscript preparation.

References

Abdul Moomin, Abdul-Aziz. 2015. Aspects of Waali phonology. Masters Thesis, University of Ghana.
Akinlabi, Akinbiyi. 2009. Neutral vowels in Lokaa harmony. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 54(2). 197–228.
Aksënova, Alëna, Jonathan Rawski, Thomas Graf & Jeffrey Heinz. 2024. The computational power of harmonic forms. In Harry van der Hulst & Nancy A. Ritter (eds.), *The Oxford handbook of vowel harmony*, 437–451. Oxford: Oxford University Press.
Angsongna, Alexander. 2023. *Aspects of the morphophonology of Dagaare*. Vancouver: University of British Columbia.
Angsongna, Alexander & Samuel Akinbo. 2022. Dagaare (Central). *Journal of the International Phonetic Association* 52(2). 341–367.
Angsongna, Alexander & Adams Bodomo. forthcoming. Language and biodiversity in context: Rethinking Dagaare dialects through indigenous plant nomenclature, PREPRINT (Version 1) available at Research Square https://doi.org/10.21203/rs.3.rs-6873791/v1 (accessed 12 August 2025).
Anttila, Arto & Adams Bodomo. 1996. *Stress and tone in Dagaare*. New Brunswick: Rutgers Optimality.
Anttila, Arto & Adams Bodomo. 2009. Prosodic morphology in Dagaare. In *Selected Proceedings of the 38th Annual Conference on African Linguistics (ACAL 38)*, 56–68. Somerville: Cascadilla Proceedings Project., no. ACAL 38: 56–68.
Archangeli, Diana & Douglas Pulleyblank. 1989. Yoruba owel harmony. *Linguistic Inquiry* 20(2). 173–217.
Archangeli, Diana & Douglas Pulleyblank. 2017. Phonology as an emergent system. In *The Routledge handbook of phonological theory*, 476–503. London: Routledge.
Archangeli, Diana & Douglas Pulleyblank. 2022. *Emergent phonology*, vol. 7. Berlin: Language Science Press.
Archangeli, Diana B. & Jonathan Yip. 2019. Assamese vowels and vowel harmony. *Journal of South Asian Languages and Linguistics* 6(2). 151–183.
Bakovic, Eric. 2000. *Harmony, dominance and control*. New Brunswick: Rutgers The State University of New Jersey.

Bakovic, Eric. 2003. Vowel harmony and stem identity. *San Diego Linguistic Papers* 1. 1–42.
Bodomo, Adams B. 1994. Language, history and culture in Northern Ghana: An introduction to the Mabia Linguistic Group. *Nordic Journal of African Studies* 3(2). 25–43.
Bodomo, Adams. 1997. *The structure of Dagaare*. Stanford: CSLI publications.
Bodomo, Adams. 2020. Mabia: Its etymological genesis, geographical spread, and some salient genetic features. In *Handbook of the Mabia languages of West Africa*, 5–34. Berlin: Galda Verlag Glienicke.
Boyd, Virginia L. Ginger. 2015. *The phonological systems of the mbam languages of Cameroon with a focus on vowels and vowel harmony*. Utrecht: Netherlands Graduate School of Linguistics.
Boyd, Virginia. 2024. Vowel harmony in Bantu Niger-Congo languages. In Harry van der Hulst & Nancy A. Ritter (eds.), *The Oxford handbook of vowel harmony*, 643–657. Oxford: Oxford University Press.
Burness, Phillip Alexander, Kevin James McMullin, Jane Chandlee, Phillip Burness & Kevin McMullin. 2021. Long-distance phonological processes as tier-based strictly local functions. *Glossa: A Journal of General Linguistics* 6(1). 1–37.
Casali, Roderic F. 1995a. Labial opacity and roundness harmony in Nawuri. *Natural Language & Linguistic Theory* 13(4). 649–663.
Casali, Roderic F. 1995b. On the reduction of vowel systems in Volta-Congo. *African Languages and Cultures* 8(2). 109–121.
Casali, Roderic F. 2003. [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7(3). 307–382.
Casali, Roderic F. 2008. ATR harmony in African languages. *Linguistics and Language Compass* 2(3). 496–549.
Casali, Roderic F. 2016. Some inventory-related asymetries in the patterning of tongue root harmony systems. *Studies in African Linguistics* 45(1 & 2). 96–99.
Casali, Roderic F. 2024. Tongue-root harmony and vowel inventory structure. In Harry van der Hulst & Nancy A. Ritter (eds.), *The Oxford Handbook of Vowel Harmony*. (2024; online edn, Oxford Academic, 22 Oct. 2024). https://doi.org/10.1093/oxfordhb/9780198826804.013.15 (accessed 9 Aug 2025).
Chandlee, Jane. 2014. *Strictly local phonological processes*. Newark: University of Delaware.
Clements, George N. 1980. *Vowel harmony in nonlinear generative phonology*. Bloomington: Indiana University Linguistics Club Bloomington.
Clements, George N. 1985. The geometry of phonological features. *Phonology* 2(1). 225–252.
Dakubu, Mary Esther Kropp. 2005. *Collected language notes on Dagaare grammar*. Accra: Institute of African Studies, University of Ghana.
Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra: Ghana Universities Press.
Dundaa, Mark. G. 2013. *Aspects of Birfor phonology*. Accra: Masters Thesis, University of Ghana.
Gafos, Adamantios & Amanda Dye. 2011. Vowel harmony: Transparent and opaque vowels. *The Blackwell Companion to Phonology* 4. 2164–2189.
Goldsmith, John. 1976. Autosegmental phonology. Ph.D dissertation, MIT.
Gouskova, Maria. 2018. Morphology and phonotactics. In *Oxford research Encyclopedia of linguistics*. Oxford: Oxford University Press.
Graf, Thomas. 2017. Locality domains and phonological c-command over strings. In *Proceedings of North East Linguistic Society (NELS 48)*. Amherst: GLSA Publications.
Heinz, Jeffrey. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41(4). 623–661.
Heinz, Jeffrey. 2011. Computational phonology–part II: Grammars, learning, and the future. *Language and Linguistics Compass* 5(4). 153–168.
Heinz, Jeffrey. 2018. The computational nature of phonological generalizations. In Larry M. Hyman & Frans Plank (eds.), *Phonological typology*, 126–195. Berlin: De Gruyter Mouton.
Heinz, Jeffrey & William Idsardi. 2013. What complexity differences reveal about domains in language. *Topics in Cognitive Science* 5(1). 111–131.
Heinz, Jeffrey & Regine Lai. 2013. Vowel harmony and subsequentiality. In *Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13)*, 52–63.
Heinz, Jeffrey, Chetan Rawal & Herbert G. Tanner. 2011. Tier-based strictly local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 58–64.
Heinz, Jeffrey, Jonathan Rawski, E. Dresher & Harry van der Hulst. 2022. History of phonology: Learnability. In *The Oxford handbook of the history of phonology*. Oxford: Oxford University Press.
Jardine, Adam & Kevin McMullin. 2017. Efficient learning of tier-based strictly k-local languages. In *International Conference on Language and Automata Theory and Applications*, 64–76. Cham: Springer.
Jarosz, Gaja. 2013. Learning with hidden structure in optimality theory and harmonic grammar: Beyond robust interpretive parsing. *Phonology* 30(1). 27–71.
Jurgec, Peter. 2011. *Feature spreading 2.0: A unified theory of assimilation*. Tromso: University of Tromso.
Kabak, Barış & Irene Vogel. 2001. The phonological word and stress assignment in Turkish. *Phonology* 18(3). 315–360.
Kaun, Abigail. 1995. *The typology of rounding harmony: An optimality theoretic approach*. Los Angeles: University of California.
Kaun, Abigail. 2004. The typology of rounding harmony. *Phonetically Based Phonology* 87416. 87–116.
Kaun, Abigail & Adam G. McCollum. 2024. Rounding harmony. In van der Hulst Harry & A Ritter Nancy (eds.), *The Oxford handbook of vowel harmony*. Oxford: Oxford University Press.
Kenstowicz, Michael & Charles Kisseberth. 2023. *Generative phonology: Description and theory*. Leiden: Brill.
Kirchner, Robert. 1993. Turkish vowel harmony and disharmony: An optimality theoretic account. *Rutgers Optimality Workshop I* 22. 1–20.
Krämer, Martin. 2003. *Vowel harmony and correspondence theory*, vol. 66. Berlin: Walter de Gruyter.
Kuubezelle, Nerius. 2013. *An autosegmental analysis of phonological processes in Dagara*. Accra: University of Ghana.

Lambert, Dakotah, Jonathan Rawski & Jeffrey Heinz. 2021. Typology emerges from simplicity in representations and learning. *Journal of Language Modelling* 9. 151–194.
Mccarthy, John J. 2011. *Doing optimality theory: Applying theory to data*, vol. 27. Malden: John Wiley & Sons.
McCarthy, John J. & Alan Prince. 1993. Generalized alignment. In *Yearbook of morphology 1993*, 79–153. Dordrecht: Springer.
McCarthy, John J. & Alan Prince. 1999. Faithfulness and identity in prosodic morphology. *The Prosody-Morphology Interface* 9. 218–309.
Mwinlaaru, Isaac N. & Foong Ha Yap. 2017. A tale of two distal demonstratives in Dagaare: Reflections on directionality principles in grammaticalisation. *Language Sciences* 64. 130–151.
McMullin, Kevin James. 2016. *Tier-based locality in long-distance phonotactics: Learnability and typology*. Vancouver: University of British Columbia.
Mwinlaaru, Isaac N. 2021. Deixis in the Dagaare nominal group: Syntagmatic and paradigmatic perspectives. *Word* 67(3). 281–317.
Mwinlaaru, Isaac N. & Foong Ha Yap. 2021. Syntactic position, qualitative features and extended demonstrative functions: Dagaare distal demonstratives Nɛ̀ and Lɛ̀ in interactional discourse. *Journal of Pragmatics* 182. 265–292.
Odden, David. 1991. Vowel geometry. *Phonology* 8(2). 261–289.
Odden, David. 1994. Adjacency parameters in phonology. *Language* 7. 289–330.
O'Keefe, Michael. 2004. Akan vowel harmony. *BA Honours Thesis, Swarthmore College*.
Ozburn, Avery, Samuel Akinbo, Alexander Angsongna, Murray Schellenberg & Douglas Pulleyblank. 2018. Dagaare [a] is not neutral to ATR harmony. *The Journal of the Acoustical Society of America* 144(3). 1938.
Pearce, Mary & Joseph Lovestrand. 2024. Vowel harmony in Chadic languages. In Harry van der Hulst & Nancy A. Ritter (eds.), *The Oxford Handbook of Vowel Harmony*. 2024; online edn, Oxford Academic, 22 Oct. 2024. https://doi.org/10.1093/oxfordhb/9780198826804.013.48 (accessed 9 Aug 2025).
Prince, Alan & Peter Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. Malden: John Wiley & Sons.
Pulleyblank, Douglas. 2002. Harmony drivers: No disagreement allowed. *Annual Meeting of the Berkeley Linguistics Society* 28. 249–267.
Ritter, Nancy A. & Harry van der Hulst. 2024. *The Oxford handbook of vowel harmony*. Oxford: Oxford University Press.
Roca, Iggy. 2003. *Generative phonology*. London: Routledge.
Rose, Sharon. 2018. ATR vowel harmony: New patterns and diagnostics. *Proceedings of the Annual Meetings on Phonology* 5. 1–12.
Rose, Sharon & Rachel Walker. 2011. Harmony systems. *The Handbook of Phonological Theory* 2. 240–290.
Saanchi, James A. 2003. Aspect and the Dagaare verb. *Cahiers Voltaiques/Gur Papers* 6. 101–106.
Snider, Keith & James Roberts. 2004. SIL comparative African wordlist (SILCAWL). *Journal of West African Languages* 31(2). 73–122.
Steriade, Donca. 1981. *Parameters of metrical harmony rules*. Cambridge: Ms, MIT.
Tesar, Bruce & Paul Smolensky. 2000. *Learnability in optimality theory*. Cambridge: MIT Press.