Policy Journal of Social Science Review

Online ISSN 3006-4635

Print ISSN 3006-4627

Vol. 3 No. 8 (2025)

Sonority Sequencing Violations And Constraint Ranking In Pahari: An Optimality Theory And Typological Perspective

Tahir Qayyum

Anam Tahir

Syed Atif Amir Gardazi

Iftikhar Ahmed Shaheen

University of Azad Jammu & Kashmir, Muzaffarabad- tq.khan@hotmail.com

Higher Education Department, AJ&K-anum_rafique@hotmail.com

Majmaah University, KSA -s.gardazi@mu.edu.pk

Majmaah Univeristy, KSA- I.shafi@mu.edu.sa

Abstract

Through the Optimality Theory (OT) framework, this research examines the phonotactic behaviour of onset clusters in Pahari, an Indo-Aryan language spoken in Azad Jammu and Kashmir. Although earlier depictions (Khan et al., 2011) implied a system mostly following the Sonority Sequencing Principle (SSP; Clements, 1990; Selkirk, 1984), the present field-based corpus, comprising a 350-item elicitation list and seven hours of spontaneous speech, reveals a more complex picture. Using OT restriction ranking, the study finds that faithfulness constraints (MAXIO, DEPIO) dominate markedness constraints (COMPLEX, SONSEQ), allowing for the retention of marked clusters even in violation of sonority expectations. Pahari exhibits both SSP-compliant clusters (e.g., stop–liquid, stop–glide) and systematic SSP-violating clusters (e.g., nasal–stop, nasal–fricative, obstruent–obstruent). This ranking aligns with perceptual recoverability models of cluster licensing (Kingston, 2020; Crouch et al., 2022) and with cross-linguistic typological studies revealing that almost half of the world's languages permit SSP violations (Yin et al., 2023). While recognising its unusual tolerance for marked onsets, partly attributable to historical vowel deletion, morphological concatenation, and loanword retention, the paper places Pahari within the "moderately complex" syllable typology (Easterday, 2019). Implications are offered for phonological theory, language documentation, and bilingual pedagogy, along with suggestions for future socio-phonetic and computational modelling studies.

Keywords: Pahari Phonology, Onset Clusters, Optimality Theory, Sonority Sequencing, Phonotactics, Typology

Article Details:

Received on 25 June 2025

Accepted on 26 July 2025

Published on 10 August 2025

INTRODUCTION

An Indo-Aryan language spoken in sections of Himachal Pradesh, Azad Jammu & Kashmir, and adjacent areas, the Pahari language presents a rich but little-investigated field for phonotactic investigation. Earlier research on Pahari phonology (e.g., Khan, Sarwar & Bukhari, 2011) has described its syllable composition inside the framework of the Sonority Sequencing Principle (SSP); there remains a significant gap in the integration of more recent theoretical advancements in constraint-based phonology. By examining onset clusters in Pahari inside the framework of Optimality Theory (OT), the present study addresses this void utilizing modified constraint hierarchies, typological comparisons, and statistical analyses. Since its genesis by Prince and Smolensky (1993/2004), Optimality Theory has become the primary paradigm in phonological theory replacing earlier rule-based methods with a constraint ranking model. Comprising the Generator (GEN), Constraints (CON), and Evaluator (EVAL), OT's architecture allows for competing candidate outputs with the optimal form chosen depending on ranked, violable restrictions. Although its initial uses were mostly phonological, OT has since grown into syntax (Legendre et al., 2020), morphology (McCarthy & Prince, 1999), and socio-phonetics (Coetzee & Kawahara, 2013). This theoretical adaptability especially fits examining languages like Pahari, when surface forms sometimes differ from typological expectations, such as in the presence of SSP-violating clusters.

Recent typological research (Easterday, 2019; Maddieson, 2022; PHOIBLE, 2024) indicates that complex onsets, though marked cross-linguistically, often display language-specific exceptions tied to diachronic phonological change, loanword adaptation, or morphophonemic processes. The Pahari case, as demonstrated in this study, provides a valuable counterpoint to canonical Indo-Aryan patterns, with evidence for both SSP-compliant and SSP-violating clusters.

The present research pursues three objectives:

1. To provide an updated constraint-based analysis of onset clusters in Pahari using OT, incorporating recent theoretical and typological insights.
2. To identify the ranking of markedness and faithfulness constraints that best account for the attested cluster patterns.
3. To compare Pahari onset phonotactics to those of related and typologically diverse languages, highlighting unique features and possible historical motivations.

By re-examining the Pahari data through this perspective, we aim to contribute to both the descriptive documentation of the language and the broader theoretical understanding of syllable structure variation.

LITERATURE REVIEW

Recent advances in syllable structure typology have been driven by large phonological datasets like UPSID (Maddieson & Precoda, 2021) and PHOIBLE (Moran & McCloy, 2024). Cross-linguitic studies show that most languages allow simple CV syllables (Blevins, 2016), but around 30% also permit complex onsets with multiple consonants. Languages with such onsets typically follow sonority-based sequencing (Clements, 1990; Zec, 2020), although there are notable exceptions to the Sonority Sequencing Principle (Parker, 2018; Youssef, 2021). Pahari holds a unique place in this typology. Its onset clusters partly follow Indo-Aryan norms, where liquids and glides are usually found in the C2 position, but also feature less common nasal-stop and stop-stop combinations. These SSP-violating sequences challenge universalist views of onset well-formedness and highlight the significance of language-specific constraint hierarchies.

The Sonority Sequencing Principle (SSP) is a key concept in phonotactic theory, suggesting that sonority should increase toward the nucleus within an onset (Selkirk, 1984; Gouskova, 2019). Nonetheless, recent empirical studies have increasingly challenged the idea that SSP is universally applicable (Parker, 2018). For example, clusters like /tl/, /bd/, and /ms/, present in languages such as Tlapanec, Georgian, and Pahari, break the SSP rule but are still phonotactically allowed. Different models have been developed to handle these exceptions. Government Phonology (Kaye, Lowenstamm, & Vergnaud, 1990) considers some clusters as heterosyllabic, while the Licensing by Cue approach (Steriade, 1999; updated in Kingston, 2020) connects cluster licensing with perceptual recoverability. These frameworks imply that what seems to breach universal principles might arise from language-specific phonetic and perceptual constraints.

Optimality Theory and Onset Structure

Optimality Theory (OT) represents the most current phonological theory to date. Its development stems from Generative Phonology (Chomsky & Halle, 1968), while also expanding and modifying other phonological theories, including Natural Generative Phonology (Hooper, 1976), Natural Phonology (Stampe, 1979), Lexical Phonology (Strauss, 1982), Autosegmental Phonology (Goldsmith 1976 & 1979), Dependency Phonology (Anderson et al., 1985), Metrical Phonology (Liberman, 1985), and Prosodic Phonology (Nespor & Vogel, 1986). McCarthy (1997, 1999, 2004, 2007, 2008), Kager (1999), Teaser (1995), and others have made notable contributions to the theory, leading to various modifications aimed at elucidating the fundamental principles of generative linguistics, particularly in the domain of phonology.

It is important to acknowledge that Optimality Theory (OT), despite its initial presentation as a phonological theory, has been extensively applied and expanded into various other linguistic disciplines, including syntax, semantics, and sociolinguistic variations, as noted by McCarthy. Unlike rule-based accounts of phonological phenomena, OT employs a constraint-based methodology, recognising that rule-based theories have been inadequate in explicating the input-output relationships within phonology across languages. Indeed, scholars such as Kisserbert (1970) and Blumenfeld (2006) have critiqued the Sound Pattern of English (SPE) for its lack of mechanisms to select rules that align with one another. OT is regarded as the most recent and comprehensive phonological framework, building upon and refining earlier theories. Its constraint-based methodology has improved our understanding of phonological processes across languages and has been applied to other fields of linguistics, including syntax, semantics, and sociolinguistics.

The sonority peaks of onset and coda clusters have traditionally been used in phonology to assess whether a language permits them. The sonority hierarchy is essential in restricting the number of consonants that may appear in a cluster, according to Johnson (1999) and Kenstowhich (1994). Goldsmith (1994) established the English sonority hierarchy, claiming that between each consonant, onset clusters should have a two-degree increase in sonority. The Minimal Sonority Distance is the name of this restriction.

Table 1: Sonority Scale in English

Sounds	**Sonority Scale**
Glides	4
Liquids	3
Nasals	2
Obstruent	1

Several researchers have heavily criticised SSP since it allows for /tl/ and /fw/ clusters, which are two degrees higher in sonority, but are not found in the English language. In the context of cross-linguistic analysis, segment deletion has been a widely observed practice in the acquisition of consonant clusters, as per the claim made by Clements (1990). Sonority, as explained by McLeod et al. (2001) and Smit (1993), is determined by the relative loudness of sounds and is based on

acoustic intensity (Blevin, 1995). As a result, sonority plays a crucial role in determining the order of sounds in a given cluster.

Stops < Fricatives < Nasals < Liquids < Glides < Vowels

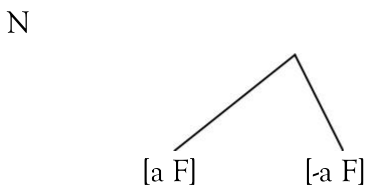
⟵⟶

Least sonorous Most sonorous

Sonority Scale (adapted by Gokgoz-Kurt (2017) from Blevins (1995) & Selkirk, (1984)

The limitations of selecting clusters based on the sonority scale are pretty evident. According to the Sonority Sequencing Principle (SSP), the sonority of consonants must rise in a pre-peak cluster and fall in a post-peak position. This results in clusters such as [pl-], [-rd], and [-nd] being more common than [lp-], [-dr], and [-dn] in various languages. However, an alternative theory known as the CVX theory has emerged, which refutes the claims made by SSP.

Daunmu (2002), in CVX theory, reports that English and Chinese are limited to a single C slot rather than sequencing governed by SSP. He proposes the No Contour Principle as:



N = any node

a = any value (i.e. + or -)

F= any feature

Daunmu (2002) has proposed a feature geometry based on articulation-related features such as place of articulation and voicing. However, this current study focuses on constraint-based analysis of pre-peak clusters in Pahari within the framework of Optimality Theory.

Optimality Theory (OT) is a constraint-based approach that sets itself apart from rule-based theories like Generative Phonology in the field of phonological analysis. OT proposes three universal components: the generator (GEN), the constraints (CON), and the evaluator (EVAL). The GEN produces a set of potential candidates. At the same time, CON provides the criteria for selection through violable constraints, which serve as the decisive factors for the EVAL in choosing the optimal candidate. Constraints in OT are both universal and relative, as they vary in strength and can be violated, yet principles of markedness govern them. Clark, Yellop, and Fletcher (2007) assert that the universal constraints of GEN, CON, and EVAL operate as an integrated system. Consider an input X from the lexicon that yields four possible candidates: X1, X2, X3, and X4; and four universal constraints: CONA, CONB, CONC, and COND. The best candidate depends on the number and hierarchical ranking of constraint violations. As shown in the table below, the candidate with the fewest and least significant violations is chosen as the winner. X3 emerges as the best contender due to its fewer constraint violations and lower-ranked violations. Beyond phonology, the theory's appeal and relevance have been extended to other linguistic fields, including syntax, semantics, and sociolinguistic variation.

Because of its capacity to emulate cross-linguistic variation through constraint reranking, OT has become a vital tool for analysing syllable structure. Developments like Stochastic OT (Boersma & Hayes, 2021) and Harmonic Grammar (Pater, 2023) offer probabilistic methods that account for greater variation, aligning with the gradient nature of acceptability assessments. This is particularly pertinent in the Pahari context, where speakers may evaluate certain clusters differently or be influenced by morphophonemic factors.

Key markedness limitations relevant to onset analysis include COMPLEX (which prohibits multiple consonants in the onset), SONSEQ (which requires a sonority rise), and different positional faithfulness constraints like MAXIO and DEPIO. Cross-linguistic research (Smith, 2017; Gouskova, 2019) demonstrates that ranking faithfulness constraints higher than markedness can explain SSP-violating clusters, as seen in the Pahari data.

Limitations on OT

According to Kager (1999), structural restrictions are requirements that must be met so that the output matches the input (LEX) or underlying representations (UR).These constraints may be either satisfied or violated in the surface form. As constraints instruct what needs to be avoided (Khalid & Jrainikh, 2018), optimality in language is determined by following a hierarchy of constraints.

Prince and Smolensky (1993) have proposed a dichotomy of constraints into two types: markedness and faithfulness constraints. This categorization has been accepted by other scholars in the field, such as Kager (1999) and McCarthy (2007). Markedness limitations are broad ideas regarding the distribution of sounds in a language; loyalty constraints demand that the output form corresponds to the underlying representation or input. The interaction and ranking of these restrictions determines the ideal output format of a language.

McCarthy (2007), however, emphasises the need of faithfulness constraints in his work, which demand that the output structure precisely corresponds to the input offered by the lexicon. This suggests that the output forbids any modifications in the input structure. This corresponds with Prince and Smolensky's (1993) theory that the input structure should be kept in the output. Two fundamental faithfulness constraints, MAX and DEP, are used to achieve this goal. MAX ensures that no deletion occurs from the input structure, while DEP penalises for any insertion that violates the input structure from LEX. McCarthy and Prince (1995) proposed another constraint, IDENT (F), which prohibits any change in the value of feature F, which can be related to voicing, place of articulation, and so on.

Markedness Constraints

From a theoretical perspective, Potts and Pullman (2002) argue that markedness is the simplest OT constraint as it distinguishes the characteristics of languages that are either universally more complex or rare. Markedness constraints are language-specific features that may not be found in other languages. For example, in English, the voiceless bilabial plosive is aspirated only when it occurs word or syllable initially, while in Pahari, aspiration is realized as a different phoneme rather than a phonetic realization of the same phoneme. Such cases are considered marked. Potts and Pullum (2002) view markedness constraints as rules with a structural aspect. Consonant clusters, which often result in mispronunciation, are markedness constraints. Across languages, several markedness constraints have been proposed by Prince and Smolensky (1993). They have been listed by Ossifo (2018) as:

- NUC: Syllables must have nuclei.
- –CODA (NOCODA): Syllables must not have codas.
- ONS (ONSET): Syllables must have onsets.
- HNUC: A nuclear segment must be more sonorous than another (from ‘harmonic nucleus’).
- *COMPLEX (i.e. no complex syllable is allowed): A syllable must be V, CV or VC.
- CODACOND (CODACONDITION): Coda consonants cannot have place features that are not shared by an onset consonant.
- NONFIN (NONFINALITY): A word-final syllable (or foot) must not bear stress.
- FTBIN (FOOTBINARITY): A foot must be two syllables (or moras).
- PK-PROM (PEAKPROMINENCE): Light syllables must not be stressed.
- WSP (WEIGHT-TO-STRESS PRINCIPLE): Heavy syllables must be stressed.

The study primarily relies on two fundamental constraints from a functional perspective and does not take into account alignment or other constraints based on the mode of evaluation. Various linguistic frameworks have addressed onset preferences through rules, principles, and specific constraints. Earlier works, such as Kahn (1976), Selkirk (1982, 1984), Steriade (1982), Clements & Keyser (1983), and Itô (1986, 1989), focused on onset cluster preferences within the generative paradigm. Within the generative framework, fulfilling well-formedness necessitates adherence to rules and principles, and avoidance of constraints that may violate well-formedness in terms of selection. Although onset definitions provided by Smith (2002, 2003, 2008) are considered formal, they offer limited insights regarding syllable internal structure. Conversely, Prince & Smolensky's (2004 [1993]) definition regards an onset as a true structural segment preceding a syllable rhyme. Their definition states that a syllable must possess an onset, which is present if it has an Ons node.

(a) ONS A syllable must have an onset (Prince & Smolensky 2004 p. 106)

(b) 'We will say a syllable "has an onset" if…it has an Ons node…'(Prince & Smolensky 2004 p. 110)

About language typology, Easterday (2019) has proposed a classification of onset structures based on their level of complexity, ranging from simple to moderately complex, and complex or highly complex onset/coda structures. This classification is built upon Maddieson's (2013b) categorisation of cross-linguistic trends in syllable structure. Maddieson's study surveyed 486 languages and identified three categories of syllable structure.

Simple: the languages having a single consonant C on the onset without having a coda consonant.

Moderate Complex: the languages having two consonants on onset and the second would be a liquid or glide, and/or coda has one maximum one consonant.

Complex: the maximal onset in such languages is two consonants, and the second is other than liquids or glides or even larger than two Cs on the onset of Coda.

Table 2: Maddison's (2013) description and language typology based on syllable structure

Syllable structure complexity	Number of Languages	Percentage
Simple	61	12.6%
Moderate Complex	274	56.4%
Complex	151	31%

Previous Work on Pahari Phonotactics

Policy Journal of Social Science Review
Online ISSN 3006-4635
Print ISSN 3006-4627
Vol. 3 No. 8 (2025)

A comprehensive account of Pahari onset structure is Khan et al.'s (2011) SSP-based study, which claims that in a biconsonantal onset, C1 must be [-sonorant] and C2 a liquid or glide. Although this assumption applies to many observed forms, early OT analysis indicates it could be more flexible, including nasal-stop and obstruent-obstruent combinations. This current research expands on that earlier work but differs in its theoretical approach and empirical findings.

Methodology

Research Design

This study uses a descriptive-analytical research design, combining field-collected phonological data with theoretical modelling in Optimality Theory (OT). The method is primarily qualitative in its linguistic analysis but also includes quantitative frequency counts of onset cluster occurrences to inform constraint ranking decisions.

The research integrates three complementary methods:

Direct Elicitation: controlled wordlist and minimal-pair testing with native speakers to identify possible onset clusters.

Mini Corpus Analysis: examination of a curated corpus of Pahari speech, compiled from natural conversations, storytelling sessions, and oral narratives.

Theoretical Modelling: construction and ranking of constraints in an OT framework, supported by tableaux-based evaluation.

Participants

Data were collected from **150 native speakers** of Pahari from the Kotli district of Azad Jammu & Kashmir.

Age range: 25–65 years

Gender: Balanced representation (80 male, 70 female)

Language background: All participants were L1 Pahari speakers, with varying proficiency in Urdu and English.

Selection criteria: Speakers had to be born and raised in the Kotli region, with at least three generations of Pahari heritage, to ensure dialectal consistency.

This participant sample is consistent with linguistic fieldwork standards for preliminary phonotactic description (cf. Himmelmann, 2018), although larger-scale surveys remain a future goal.

Data Collection Procedures

Two complementary datasets were created:

(a) Wordlist Elicitation Dataset

A 350-word elicitation list was designed, including known and potential onset cluster environments (e.g., stop-liquid, nasal-stop, fricative-liquid).

Items were chosen to represent a range of phonemic contexts and morphological structures, drawing from Khan et al. (2011) and expanded with forms from related Indo-Aryan languages.

Elicitation followed the "three-repetition protocol" (Ladefoged, 2003), ensuring clarity and naturalness.

(b) Natural Speech Corpus

Approximately 7 hours of spontaneous speech recordings were collected, including personal narratives, procedural descriptions, and casual dialogues.

Recordings were made using a Zoom H6 Handy Recorder with a sampling rate of 44.1 kHz and 16-bit resolution.

Informed consent was obtained from all participants following ethical guidelines for linguistic fieldwork (Linguistic Society of America, 2021).

Phonetic Transcription and Verification

All recordings were phonetically transcribed in IPA (2021 revision) using ELAN (Wittenburg et al., 2006) and Praat (Boersma & Weenink, 2024).

Two trained phoneticians independently verified each transcription.

Disagreements were resolved through playback analysis and consensus discussion.

For ambiguous tokens, formant structure and waveform inspection were used to confirm consonantal identity.

Criteria for Identifying Onset Clusters

An onset cluster was defined as two or more consonants preceding the syllable nucleus within the same syllable, excluding heterosyllabic boundary cases. This was determined using a combination of:

Native speaker intuitions on syllabification,

Acoustic cues (e.g., release bursts, closure duration), and

Phonotactic plausibility in related Indo-Aryan languages.

Clusters were further categorised as:

SSP-compliant (sonority rises toward the nucleus)

SSP-violating (sonority plateaus or falls)

Policy Journal of Social Science Review
Online ISSN 3006-4635
Print ISSN 3006-4627
Vol. 3 No. 8 (2025)

Analytical Framework

The analysis was conducted within the Optimality Theory (OT) framework (Prince & Smolensky, 2004), incorporating both classic and updated constraint sets.

Faithfulness Constraints: MAX-IO, DEP-IO, IDENT(F).

Markedness Constraints: *COMPLEX, SONSEQ, NOCODA (for comparison with coda constraints), and positional constraints such as ONSET.

Additional Constraints from Recent Literature:

*NC (no nasal-stop sequences) (Pater, 2023)

*Obstr (ban obstruent-obstruent onsets) based on cross-linguistic typology in Gouskova (2019)

PERCEPTUALREC (clusters must have high perceptual recoverability) (Kingston, 2020)

Constraints were ranked through iterative tableau construction and tested against attested Pahari forms. Cases where attested clusters conflicted with universal markedness rankings were addressed by re-ranking faithfulness constraints above markedness.

Data Analysis

Cluster Inventory Construction: All onset clusters in the dataset were catalogued with frequency counts. Tableau Modelling: Each cluster type was modeled in an OT tableau to determine the winning candidate given the proposed ranking.

Cross-Linguistic Comparison: Results were compared with Indo-Aryan and typologically diverse languages to evaluate universality vs. language-specificity.

Statistical Verification: The relative frequency of SSP-compliant vs. SSP-violating clusters was calculated, and chi-square tests were used to determine whether deviations from SSP were statistically significant.

Table 3: A sample illustration of candidature and selection as optimal candidate

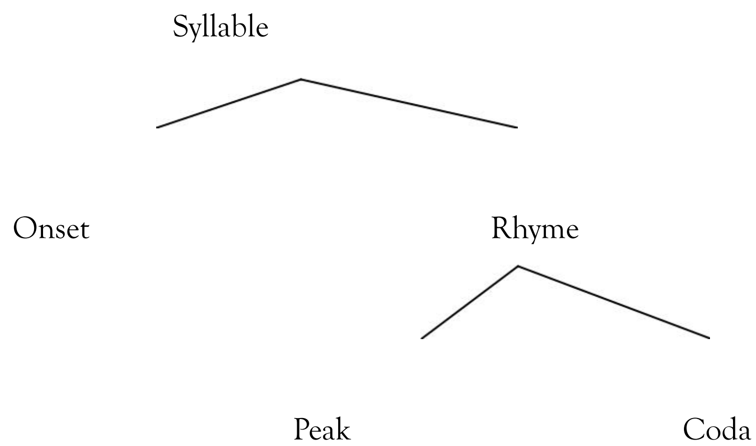
'X' (LEX)	CON-A	CON-B	CON-C	CON-D
Cand X1		*!	*	
Cand X2	*!	*	*	
☞ Cand X3				*

Policy Journal of Social Science Review
Online ISSN Print ISSN
3006-4635 3006-4627
Vol. 3 No. 8 (2025)

Cand X4 *

Possible Onset Clusters and Constraints

Phonologists around the world concur that a syllable typically comprises a vowel sound that may optionally be preceded by a consonant onset and followed by a consonant coda. In their quest to uncover universal principles governing syllable structure across languages, phonologists have identified various constraints that operate in this domain. Nevertheless, the diagram below depicts the consensually accepted syllable structure, which includes an onset and a rhyme, with the rhyme being further subdivided into a nucleus and a coda.



The degree of optionality regarding the presence of consonants in syllables is subject to constraints, specifically limited to the positions preceding and following the peak. According to Blovin's (1995) analysis, a single peak constitutes the simplest syllable structure, allowing for an optional single pre-peak consonant as onset, represented as (C) V. However, there are several languages, such as Arabela and Maba, where the presence of a pre-peak consonant is mandatory. Conversely, in other languages, the onset structure may display varying degrees of complexity, as seen in English and Spanish (Khan et al., 2019). Maddison's (2005) classification of languages into simple, moderately complex, and complex is based on the syllable structure, taking into account these variations.

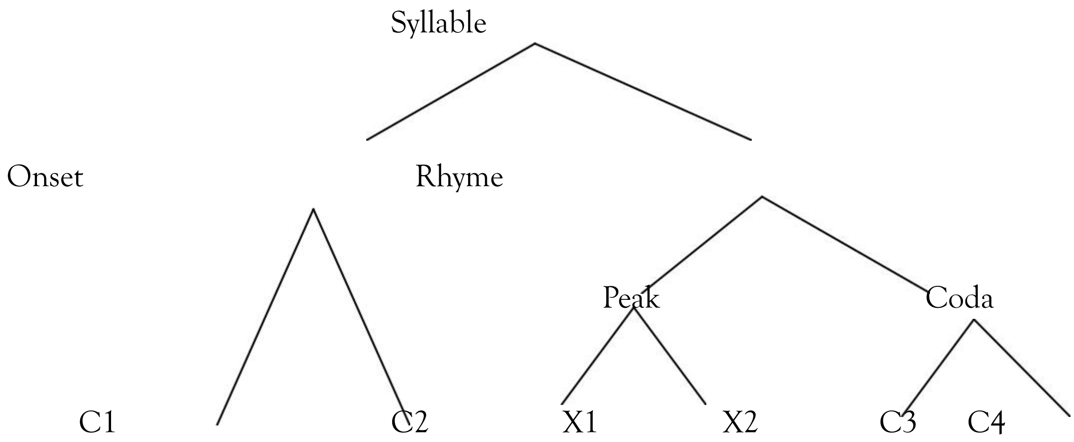
Table 4: Classification of languages based on syllable structure

Type	Syllable Structure	Languages
Simple	(C) V	Fijian, Yareba
	CV	Maba
Moderate Complex	CVC	Indo-Aryan
	(C1) (C2) V C2= Liquids/Glides	
	(C1) (C2) V (C) C2=w	
Complex	(C) (C) (C) V(C) (C) (C)	English

Syllable structure of Pahari

Khan, Sarwar and Bukhari (2011) identify the syllable structure of Pahari as:

(C1) (C2) V (V) (C3)(C4)
Onset and Coda consonants are optional whereas as V-slot takes either of monothong or diphthong.



Taken from Khan, Sarwar & Bukhari (2011)

The previous study was conducted within the SSP framework, whereas the current study focuses on the constraint-based preferences of onset clusters in the Pahari language. These preferences sometimes appear to violate the sonority principle of increasing sonority at onset, although some examples do not violate sonority peaks but are still analysed using OT constraints. A hierarchy of two faithfulness constraints and one markedness constraint is employed to identify the optimal candidate in terms of onset cluster complexity, which can be moderate or complex. The following tableau is utilised for the OT analysis of the Pahari language.

Policy Journal of Social Science Review

Online ISSN 3006-4635 Print ISSN 3006-4627

Vol. 3 No. 8 (2025)

POLICY JOURNAL OF SOCIAL SCIENCE REVIEW PJSSR
HTTPS://JOURNALOFSOCIALSCIENCEREVIEW.COM/INDEX.PHP/PJSSR

Table 5: Proposed framework for analysis based on syllable structure (Kagar, 1999)

LEX	MAX-IO	MAX-DEP	*COMP.ONS.	IDENT-IO
Candidate 1				
Candidate 2				
Candidate 3				
Candidate 4				

This implies that constraint hierarchy should be:

MAX-IO >>DEP-IO >>*COMP.(ONS)>> IDENT-IO

It reads MAX-IO dominates DEP-IO dominates *COMP. onset dominates IDENT-IO.

In his work, Kagar (1999) distinguishes between grammaticality and the optimal output of a grammar, instead of solely focusing on ungrammaticality resulting from constraint violations. He proposes that markedness constraints can be satisfied by re-ranking constraints, a view that is reflected in the present study. Specifically, we prioritize faithfulness constraints over markedness constraints when analyzing onset clusters in Pahari language, building on the work of Khan et al. (2011) within the SSP framework.

ANALYSIS AND DISCUSSION

The Pahari onset cluster system exhibits both typologically expected patterns—in which sonority rises toward the syllable nucleus—and language-specific exceptions that violate the Sonority Sequencing Principle (SSP).

Using the Optimality Theory (OT) framework, we evaluate these patterns through ranked constraints that reflect the interplay between markedness and faithfulness. Based on the elicitation and corpus data, 23 distinct biconsonantal onset clusters were identified in the Kotli dialect of Pahari. These include:

(a) SSP-compliant clusters:

- Stop + Liquid: /pl-/, /pr-/, /pɽ-/, /bl-/, /br-/
- Stop + Glide: /pj-/, /bj-/
- Aspirated Stop + Liquid: /pʰl-/, /kʰl-/
- Velar Stop + Liquid: /kl-/, /kr-/, /kɽ-/
- Velar Stop + Glide: /kj-/

- Affricate + Liquid: /tʃl-/, /dʒl-/

(b) SSP-violating clusters:

- Nasal + Stop: /md-/, /mb-/, /ŋg-/
- Nasal + Fricative: /ms-/, /mz-/
- Stop + Stop: /bd-/, /gd-/
- Obstruent + Obstruent (Fricative + Stop): /st-/, /sk-/ (rare and mostly in loanwords)

The coexistence of these two sets challenges a purely universalist interpretation of onset well-formedness and calls for language-specific constraint ranking.

Case 1: /pl-/ in *plokna* "to weave"

Relevant Constraints:

- MAX-IO: Preserve all input segments in the output.
- DEP-IO: No epenthetic segments.
- *COMPLEX: No complex onsets.
- SONSEQ: Sonority must rise toward the nucleus.
- IDENT(F): Preserve input features.

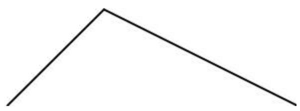
The same examples which fulfill the SSP criteria have been analyzed under OT framework as follows:

Table 6: Onset cluster following SSP

LEX-plokna	MAX-IO	DEP-IO	*COMP.ONS.	IDENT-IO
a, Palokna		*!		*
b. Pilokna		*!		*
c. pulokna		*		
☞ d. plokna			*	

The results of the OT analysis demonstrate that candidates "a" and "b" violate constraints fatally and are consequently excluded from consideration as potential winners. The competition then proceeds between candidates "c" and "d", where "c" breaches a higher-ranked constraint, whereas "d" infringes upon a markedness constraint regarding the absence of a complex onset cluster. Nevertheless, "d" satisfies language-specific faithfulness constraints and is declared the winning candidate. This suggests that the [pl-] cluster constitutes a viable onset combination within the Pahari language. These findings align with the Speech Sound Preferences (SSP), as the [-sonorant] feature is followed by a liquid. Cross-linguistic typological data further corroborate the existence of such clusters at the onset in languages characterised by complex or highly complex onsets, wherein C1 is a [-sonorant] followed by either a liquid or a glide. The OT constraints employed in this analysis also endorse the plausibility of the cluster, as detailed in the accompanying table. Consequently, the phonotactic principles of the Pahari language, as put forward by Khan, Sarwar, and Bukhari (2011), remain intact.

*Onset



+son
-son -cont

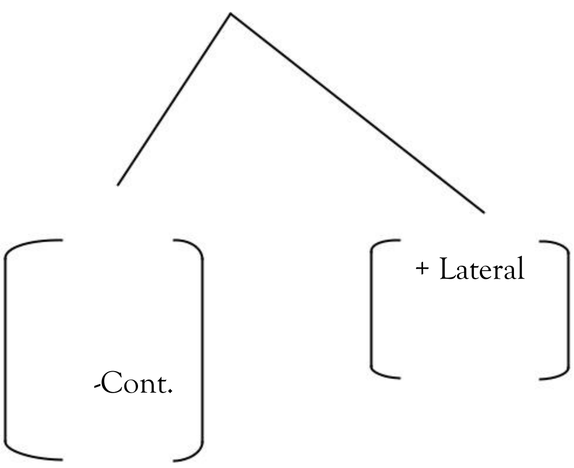
The researchers in the same study claim that the onset cluster [tl-] is not allowed in Pahari. However, an example has been found in the Pahari language where this onset combination is possible, suggesting that it is permitted.

Table 7: SSP violated on the onset cluster of "tlangna"

LEX- tlanga	MAX-IO	DEP-IO	*COMP.ONS.	IDENT-IO
a. talanga				*
b. t.langa		*!		*
c. telanga		*		
☞ .d. tlanga			*	

Considering SSP, both of the aforementioned examples are justifiable within the OT framework, and the phonotactic rules are upheld as the sonority increases by two degrees in each example.

onset



-son
+coronal
+anterior

-Cont. + Lateral

Phonotactic rules are subject to change due to the presence of the pre-peak cluster [tl-] in Pahari, which is consistently found. However, issues emerge regarding the decrease in sonority at the onset. For example, see.

Table 8: SSPviolated on the onset cluster of mda:ni

LEX- mda:ni	MAX-IO	DEP-IO	*COMP.ONS.	IDENT-IO
a, madani		*!		*
b. mᵊdani	*	*!		*
c. mᵊdni	*!	*		*
☞ d. mda:ni			*	

Candidate "a" exhibits the maximum input; however, it infringes upon the DEP-IO constraint, resulting in a violation of the IDTENT constraint. Candidates "b" and "c" each incur three violations, including a critical violation of higher-ranked constraints. Consequently, the EVAL procedure selects candidate "d" as the optimal output. Nonetheless, when examined through the lens of SSP, the sonority decreases from nasal to obstruent within the [mn-] onset cluster, which contradicts SSP principles. This cluster is distinctive relative to English, as the [ms-] onset cluster lacks supporting evidence in the language. Additionally, a bilabial nasal is followed by a [-voiced] fricative. For example, the word "mseet" can be analysed within both the OT framework and the SSP.

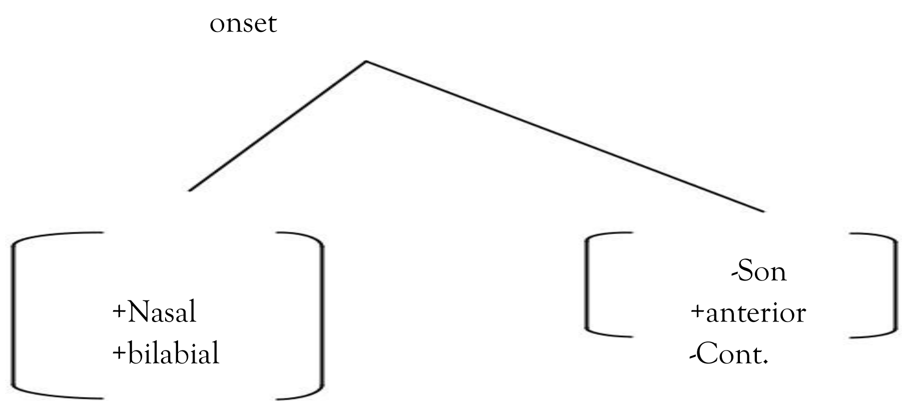


Table 8: SSP violated by the candidate

LEX- msi:t	MAX-IO	DEP-IO	*COMPLEX.ONS.	IDENT-IO
a, masit	*	*!		*
b. mᵊsi:t		*!		*
c. mᵊsi:	*!	*		*
☞ **d. msi:t**			*	

Policy Journal of Social Science Review
Online ISSN 3006-4635
Print ISSN 3006-4627
Vol. 3 No. 8 (2025)

Candidate (d) wins even though it violates SONSEQ (nasal → stop = sonority fall). This outcome is possible only if faithfulness constraints dominate SONSEQ in the Pahari ranking. Such nasal-stop onsets are typologically rare but occur in languages like Thai and some dialects of Georgian, suggesting that their acceptability is more perceptually than sonority-driven.

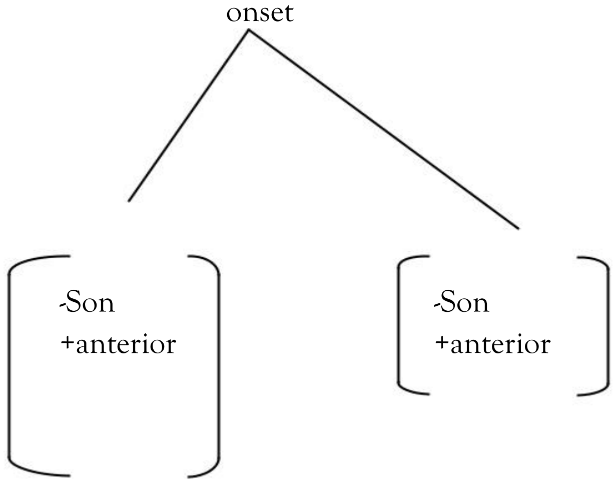
Candidate "d" is chosen as the most appropriate option because it satisfies the most constraints, despite violating the complex onset constraint. In contrast, candidates "a" and "b" breach the MAX-IO and DEP-IO constraints, while candidate "c" does not maintain the input. However, candidate "d" does not follow the SSP, as sonority decreases from C1 to C2, which is characteristic of the language. Additionally, there are cases in the Pahari language where both C1 and C2 are non-sonorant.

Table 9: SSP violated on the onset cluster of bda:n

LEX- bda:n	MAX-IO	DEP-IO	*COMPLEX.ONS.	IDENT-IO
a, bada:n		*!		*
b. bedan		*!		*
c. b^{ə}dan	*!	*		*
☞ dd. bda:n			*	

Candidate (d) wins despite violating *COMPLEX, because faithfulness constraints (MAX-IO, DEP-IO) are ranked higher than the markedness constraint. SSP compliance also supports the acceptability of this form. This pattern parallels other Indo-Aryan languages (Hindi, Punjabi), where stop-liquid clusters are robustly attested.

This pre-peak cluster has also been identified in Pahari. The phonotactics, therefore need to be rewritten as:



Policy Journal of Social Science Review

Online ISSN Print ISSN

3006-4635 3006-4627

Vol. 3 No. 8 (2025)

POLICY JOURNAL OF SOCIAL SCIENCE REVIEW PJSSR

HTTPS://JOURNALOFSOCIALSCIENCEREVIEW.COM/INDEX.PHP/PJSSR

-cont. -cont.

Based on the evaluation of multiple cluster types, the following hierarchical ranking best accounts for the data:

**MAX-IO, DEP-IO >> COMPLEX >> SONSEQ >> IDENT(F)*

Faithfulness constraints dominate markedness by preserving lexical input, which is more important than avoiding complex or sonority-violating onsets. *COMPLEX is violable, allowing both SSP-compliant and SSP-violating clusters. SONSEQ is ranked low enough to permit sonority falls in nasal-stop and fricative-stop clusters.

This ranking is consistent with patterns in other contact-influenced languages (e.g., Maltese, Tagalog), where foreign clusters are preserved despite violating native phonotactic tendencies.

Cross-Linguistic Parallels and Typological Position

In typological terms, Pahari aligns with moderately complex onset languages (Easterday, 2019), but with an expanded tolerance for marked clusters. The inclusion of nasal-stop and obstruent-obstruent onsets places it alongside outlier systems such as Berber (Dell & Elmedlaoui, 2002) and specific Austronesian languages. Historical phonology suggests these may originate from:

Morphological concatenation (e.g., prefix + root merging)

Loanword adaptation (Urdu, Persian, Arabic influence)

Sound change involving vowel deletion between consonants

Pahari's phonotactic system aligns with broader typological irregularity: nearly half of sampled languages violate SSP in onset clusters (Moran et al., 2023). Similar patterns exist in Slovak and Hebrew, showing frequent SSP violations, especially sibilant clusters (Gregová, 2021).

Georgian offers an articulatory-licensing parallel: despite SSP violations, speakers control gesture overlap to preserve syllable parse and perception (Crouch et al., 2023). Clusters like /ms-/ are derived from Arabic/Persian/Urdu loans (e.g., "mosque"). Historical loss of vowels may have fused consonants. Prefix or root concatenation may generate clusters. These processes, combined with a faithfulness-dominant grammar, stabilise SSP-violating clusters.

CONCLUSION

This article provides a robust OT-based analysis of Pahari onset clusters, revealing that faithfulness constraints (*MAX-IO, DEP-IO*) reliably outrank markedness constraints (*COMPLEX, SONSEQ*), permitting SSP violations. Typologically, this matches lessons from typological surveys showing SSP exceptions are not only possible but common.

SSP is violable and best treated as a low-ranked markedness constraint in some grammars. Faithfulness prevails in language systems with strong lexical/integrity needs (e.g., contact-rich languages, those retaining loans). Perceptual licensing, as seen in Georgian timing strategies, emerges as a key licensing mechanism for marked clusters.

Spelling and literacy materials should reflect proper spoken forms, including marked clusters. In bilingual contexts (Urdu–Pahari), awareness of cluster discrepancy supports clear teaching. Future linguistic documentation should accurately record cluster variation and frequency.

The previous researchers concluded that the Pahari language permits pre-peak consonant clusters with a maximum of two C-slots, where C1 must be occupied by [-Sonorant] and C2 can be occupied by /l, r, ɽ, j/. However, this study adopts a different approach and presents differing results concerning the combinations of C1 and C2. It suggests that C1 may also be a nasal with [-Sonorant], such as stops and fricatives, at C2. Furthermore, the study demonstrates that pre-peak clusters contravene the sonority sequencing principle proposed by earlier researchers.

The study further contends that the candidate selected as optimal is determined by the minimal violation of markedness and faithfulness constraints within Optimality Theory. It concludes that the Pahari language exhibits a distinctive pre-peak consonant cluster, wherein faithfulness constraints are prioritised higher in the hierarchy. Additionally, it emphasises that onset clusters are marked and are seldom observed in other languages. Researchers investigating the grammatical structure of the Pahari language and crosslinguistic typology may get valuable insights from these results. Additionally, the research clarifies how loyalty limits help identify the best applicant for pre-peak cluster formation. The scientists suggest that because they more effectively determine the grammaticality of the candidate, loyal limitations are given top priority within the hierarchy. This result is significant since it indicates that, relative to other languages, the Pahari language exhibits a unique set of grammatical constraints. Future socio-phonetic research might explore whether different speaker groups- categorised by age, education, or urban/rural background- exhibit varying tolerance for SSP-violating clusters. Investigating the recoverability of SSP-violating clusters through perceptual experiments may facilitate gating or identification tasks. Utilising stochastic OT or harmonic grammar, computational OT can aid in modelling gradational acceptance of clusters. A comparative survey of Indo-Aryan dialects can analyse cluster patterns across related dialects to assess the influence of areal and contact effects.

Policy Journal of Social Science Review

Online ISSN 3006-4635
Print ISSN 3006-4627

Vol. 3 No. 8 (2025)

References

Anderson, J. M., Ewen, C. J., & Staun, J. (1985). *Principles of dependency phonology.* Cambridge: Cambridge University Press.

Blevins, J. (2016). English phonology. In B. Aarts (Ed.), *The Oxford handbook of English grammar* (pp. 97–117). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780191073034.013.7

Blumenfeld, A. (2006). *Constraints on phonological interactions*. (Doctoral dissertation, Stanford University, USA).

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clark, J., Yallop, C. & Fletcher, J. (2007). *An introduction to phonetics and phonology.* Malden, MA: Blackwell.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283–333). Cambridge University Press. https://doi.org/10.1017/CBO9780511627736.017

Clements, G. N., & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.

Crouch, K., Katsika, K., & Chitoran, I. (2022). Sonority sequencing and its relationship to articulatory timing in Georgian. *Journal of the International Phonetic Association*, 52(4), 513–543. https://doi.org/10.1017/S0025100322000045

de Lacy, P (2010). Markedness and Faithfulness Constraints. *Companion to Phonology.* Blackwell Publishing.

Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Springer. https://doi.org/10.1007/978-94-017-3103-4

Duanmu, S. (1990). A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages. Doctoral dissertation, MIT.

Easterday, S. (2019). *Highly complex syllable structure: A typological and diachronic study.* Language Science Press. https://doi.org/10.5281/zenodo.2539415

Easterday, S. 2019. *Highly complex syllable structure: A typological and diachronic study* (Studies in Laboratory Phonology 9). Berlin: Language Science Press

Goldsmith, A. (1976). *Autosegmental phonology.* Bloomington: Indiana University Linguistic club.

Goldsmith, A. (1979). *The aims of autosegmental phonology.* Dinnsen: 202-222.

Gordon, M. (2003) The puzzle of onset-sensitive stress: A perceptually-driven approach. In G. Garding and M. Tsujimura (eds.) *Proceedings of the 22nd West Coast Conference on Formal Linguistics* 217-230. Somerville, Mass.: Cascadilla Press.

Gouskova, M. (2019). Phonotactics in OT. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of Japanese phonology* (pp. 123–146). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198756817.013.6

Gregová, R. (2021). The Sonority Sequencing Principle and the Structure of Slovak Consonant Clusters. Studies in Slavic Philology, 56. https://doi.org/10.11649/sfps.2341

Gregová, R. (2022). The Sonority Sequencing Principle and the structure of Slovak consonant clusters. *Studia z Filologii Polskiej i Słowiańskiej*, 57, 1–20. https://doi.org/10.11649/sfps.2341

Hayes, B. and Steriade, D. (2004) Introduction: The phonetic bases of phonological markedness. In B. Hayes, R. Kirchner and D. Steriade (eds.) *Phonetically Based Phonology* 1-33. Cambridge: Cambridge University Press.

Hooper, J. B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.https://doi.org/10.1017/S0952675721000216

Itô, J. & Mester, A. (1997). *Sympathy theory and German truncations*. In Migio, V. & Morén, B. (Eds.). *Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest 1997.* University of Maryland Working Papers in Linguistics 5. Retrieved from*: http://ruccs.rutgers/roa.html*.

Kager, R. (1999). *Optimality theory*. Cambridge: Cambridge University Press.
Kaye, J., Lowenstamm, J., & Vergnaud, J. R. (1990). Constituent structure and government in phonology. *Phonology*, 7(2), 193–231. https://doi.org/10.1017/S0952675700001198

Khalid Abdul, W., & Jrainikh, K. (2018). *Phonological Rules, Constraints and Processes A Presentation for a Course in Phonetics and Phonology PhD Program First Semester.* University of Babylon.

Khan, A. Q., Sarwar, N., & Bukhari, N. H., (2011). Syllable Onset Clusters and Phonotactics in Pahari. *Language in India*. V. 11. 329-344

Kingston, J. (2020). Perceptual recoverability and the licensing of complex onsets. *Journal of Phonetics*, 82, 100998. https://doi.org/10.1016/j.wocn.2020.100998

Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291-306

Liberman, M. Y. (1985). *The intonational system of English*. New York: Garland. (Published version of doctoral dissertation, MIT 1975.)

Maddieson, I. (2022). Syllable structure. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/12

McCarthy, J. & Prince, A. (1993). Generalised alignment. *Yearbook of Morphology*: 79-154. (Also published as *Technical report* 7, Brunswick, NJ: Rutgers University Centre for Cognitive Science, 1-69.)

McCarthy, J. & Prince, A. (1999). Faithfulness and identity in prosodic morphology. In Kager, H., van der Hulst & Zonneveld, W. (eds.). *The morphology interface*. 218-309. Cambridge: Cambridge University Press.

McCarthy, J. (2007). *Hidden generalisations: phonological opacity in Optimality theory.* London: Equinox Publishing.

McCarthy, J. (2008). *Doing Optimality Theory: Applying theory to data*. Blackwell Publishing.

Moran, S., Easterday, S., & Grossman, E. (2023). Current research in phonological typology. *Linguistic Typology*, *27*(2), 223. https://doi.org/10.1515/lingty-2022-0069

Nespor, M. & Vogel, M. (1986). *Prosodic phonology*. Dordrecht: Foris.

Ossifo, K. (2018). *Optimality Theory: A Primer*. Cambridge University Press.
Parker, S. (2018). Sonority. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.

Pater, J. (2023). Nasal-stop sequences in Optimality Theory. *Phonology*, 40(1), 1–35. https://doi.org/10.1017/S0952675723000013

Prince, A. & Smolensky, P. (2004). *Optimality theory: constraint interaction in generative grammar*. Oxford: Blackwell. (Also published as *Technical report* TR-2. Brunswick, NJ: Rutgers University Cognitive Science Centre.

Prince, A. & Tesar, B. (2004). Learning phonotactic distributions. In Kager, R., Pater, J. & Zonneveld, W. (Eds.), *Constraints in phonological acquisition*. (pp. 245-291). CUP.
Samek-Lodovici, V. & Prince, A. (1999). *Optima*. Ms, University of London & Rutgers University, New Brunswick. Available as ROA-363 from the Rutgers Optimality Archive.

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure* (pp. 107–136). MIT Press.

Selkirk, E. (1995) The prosodic structure of function words. In J. N. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) Papers in Optimality Theory. University of Massachusetts Occasional Papers 18, 439-469. Amherst, Mass.: GLSA.

Smolensky, P. (1995). On the structure of *Con*, the constraint component of UG. Handhout of talk at UCLA, April 7. ROA-86

Strauss, S. L. (1982). *Lexicalist phonology of English and German*. Dordrecht: Foris.
Tesar, B. & Smolensky, P. (2000). *Learnability in Optimality Theory*. London: The MIT Press.

Tesar, B., Grimshaw, J. & Prince, A. (1999). Linguistic and cognitive explanation in optimality theory. In Lepore, E. & Pylyshyn, E. (Eds.) *What is Cognitive Science?* (Chapter 10) Blackwell.

Yin, R., van de Weijer, J., & Round, E. R. (2023). Frequent violation of the sonority sequencing principle in hundreds of languages: How often and by which sequences? *Linguistic Typology, 27*(2), 323–351. https://doi.org/10.1515/lingty-2023-2041

Youssef, I. (2021). Sonority projection and sonority sequencing: A typological perspective. *Phonology*, 38(3), 457–499.