

Information-theoretic applications to Hupa verbal morphology

Cameron Duval*

Abstract. Hupa (*Na:tinixwe Mixine:whe'*) is a Pacific Coast Dene language spoken in Hoopa Valley in Northern California. Like its Dene sisters, Hupa exhibits complex verbal morphology which has attracted decades of theoretical research. One approach that has yet to have been applied to these languages is information theory. Previous information-theoretic research into verbal morphology has uncovered a cross-linguistic trend of grouping predictive information closer together and finding morphemes that are more mutually-informative to the root closer to the root, which in turn reduces overall surprisal and is easier on memory constraints. However, these studies analyzed prominently suffixing languages of Afro-Eurasia. This project is the first application of these information-theoretic concepts to a Dene language to investigate if these approaches also apply to explain morpheme order in a low-resource, Indigenous American language with intricate, prominently-prefixing morphology. The results indicate similar findings to previous research. Hupa demonstrates a word-level linear morpheme order that, on average, orders most mutually-informative morphemes closest to the verb root compared to a randomized baseline. This morpheme order also resulted in an average surprisal that was more comparable to optimized morpheme orders than a randomized baseline. Morpheme-type mutual information, however, demonstrates the discrepancies between word- and templatic-level information content in Hupa, which exemplifies the word-level efficiency that Hupa shares with other languages despite the typological uniqueness of its morphological grammar.

Keywords. information theory; Dene (Athabaskan) languages; verbal morphology; typology; computational methods

1. Introduction.

1.1. THE HUPA LANGUAGE. Hupa (*Na:tinixwe Mixine:whe'*) is a Pacific Coast Dene (Athabaskan) language spoken in Hoopa Valley in Northern California. The exact number of speakers of Hupa is not known but is likely quite low, with the totality of this project comprising of data from the native speaker Verdena Parker. As with other Dene languages, Hupa demonstrates intricate verbal morphology, as seen in its morpheme template in Table 1. We see that the verb root is at the very right in Hupa which is then preceded by prominent prefixing. Hupa also has verbal suffixes, but they are quite limited and thus not observed in this project.

* I would like to thank the audiences and reviewers of SSILA 2023, ACOL 2023, and LSA 2024 for their insightful questions and commentary. I would also like to thank Drs. Justin Spence, Jack Hawkins, and Emily Morgan for being my comprehensive exam committee while at University of California, Davis and thus for their commentary and contributions to my final master's project, of which this study and its results take up a large portion. Lastly, large thank-yous to Dr. Michael Hahn (Saarland University) for his continuous advice and assistance on this subject and his extension of code which made this project possible, as well as to the native Hupa speaker Verdena Parker for without her expansive contributions to the Hupa Online Dictionary and Texts corpus this project would not be possible. Author: Cameron Duval, University of Alberta (cduval@ualberta.ca).

11	10	9	8	7	6	5	4	3	2	1	0
Adv./ Theme	Iter.	Plural	3 Subj.	Obj.	Theme	Adv.	Distr.	Mode	1/2 Subj.	Class	Stem

Table 1. Morphological template of Hupa verbs, adapted from Golla (1970) and Campbell (2007)¹

One of the most unique aspects of Hupa’s verbal morphology lies in the order. First and second person subject marking is ordered closer to the verb root than aspectual and modal marking, along with these subject markers taking up a different slot compared to third person subject marking. There is also a fair amount of derivation and inflection mixing in the Hupa template, where morphemes that are quite lexifying and derivational in nature, thematic and adverbial morphemes, are intermixed and separated by more productive and inflectional morphemes, such as subject and object marking and modal and aspectual prefixes. These orders are also quite notable for their derivations from other previously-observed cross-linguistic trends in morpheme order, namely aspectual and modal morphemes being closer to the root than subject marking and derivation generally being found closer to the root than inflection (Bybee 1985a; Greenberg 1963). An intricate example of this is in (1), which illustrates the intermixing of perfective and adverbial prefixes among subject, object, and plural marking.

- (1) *na:yuntehsdiltin-te:*
na:-ya:-n-ti-si-di-l-te:n-te:
ITER-PL-2.SG.OBJ-ADV-PERF-1.PL.SBJ-VAL-ROOT-FUT
‘We are going to take you back.’

The unique morphological template of Hupa, as well as Dene languages more generally, has attracted several competing theories of Dene word formation. An early theory was that of templatic morphology, which claims that word formation is a step-wise process involving different steps of derivation and inflection insertion, followed by phonological rules (Kari 1989; 1990; 1992). Hargus (1988) builds on the templatic theory using Lexical Phonology, claiming that phonological rules are instead applied within each domain of morpheme insertion. McDonough (2000) uses the bipartite model to describe minimality in the Dene verb as well as explain intricate phonological interactions in domains close to the verb root.

Potter (1996) proposes an application of the Mirror Principle, where morpheme order is determined syntactically within the lexical word from an underlying D-structure and the verb stem raises to merge with its prefixes into an order that is checked by the lexical item. Finally, Rice (2000) proposes a semantic scope framework where morpheme order in Dene languages is based on how their meanings combine in relation to the root. Out of these studies, one approach that has yet to have been applied is that of a quantitative, token-based approach based on statistical distribution of morphemes to their roots, which includes information theory.

¹ Class = classifier (valence and voice), 1/2 Subj. = 1st and 2nd person subject, Mode = mode and aspect, Distr. = distributive, Adv. = adverbials, Theme = thematic prefixes, Obj. = object, 3 Subj. = 3rd person subject, Iter. = iterative.

1.1. INFORMATION THEORY. Information theory is the study of how information is quantified, stored, and communicated. This project deals with two primary information-theoretic concepts: mutual information and surprisal. Mutual information (MI) is a measure of statistical codependency between two elements. As seen by its equation in (2), mutual information is essentially a ratio of how often two elements occur together versus how often they occur separately. In the context of linguistic data, mutual information measures how co-predictable or cooccurring two linguistic units are, whether they be phonemes, morphemes, or words. In morphology, we can roughly relate this concept to productivity. The less productive a morpheme is, the more dependent its existence is to on the stem, hence having higher MI conditional to the stem.

(2) Mutual information (MI)

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p_{(X,Y)}(X, Y) \cdot \log \left(\frac{p_{(X,Y)}(X, Y)}{p_X(X)p_Y(Y)} \right)$$

The second information-theoretic concept is that of surprisal, whose equation is in (3). Surprisal is a measurement of how expected a certain element (w_t) is given a certain context ($w_1 \dots w_{t-1}$). A less expected element, e.g. an unexpected word given a sentential context, would thus have a low conditional probability and in turn a higher surprisal.

(3) Surprisal

$$\text{surprisal}(w_t) = -\log_2(p(w_t | w_1 \dots w_{t-1}))$$

Mutual information and surprisal have played several roles in linguistic research over the decades. Some of mutual information's first applications come from collocate extraction using corpus data (Church & Hanks 1990; Kita et al. 1994). Cross-linguistic corpus data has exemplified a tendency to group mutually-informative words closer together, part of a concept known as information locality, which has been observed in free-word as well as head-dependent pair corpora (Futrell et al. 2019; Hahn et al. 2021a; Pothos & Juola 2007). Surprisal theory, proposed by Hale (2001), is a framework that correlates surprisal ratings to measures of sentence processing costs. Psycholinguistic experiments have observed correlations between surprisal and reading times (Monsalve et al. 2012; Smith & Levy 2008) as well as pupil size (Frank & Thompson 2012).

Hahn et al. (2021b) utilized information theory to examine cross-linguistic morpheme order tendencies in the agglutinative languages Korean, Turkish, Japanese, Sesotho, Finnish, and Hungarian, specifically a verb affix ordering tendency observed by Bybee in (4) (1985a). The study utilized corpora that ranged from 2,735 (Hungarian) to 109,323 (Korean) verbal tokens.

(4) ROOT VALENCE VOICE TAM² PERSON³

This study observed, as seen in Figure 1 (left), that this morpheme order correlated with ordering morpheme types that are most mutually-informative to the verb root closer to the verb root. In other words, morpheme types that were more statistically co-dependent to the verb root,

² Tense, aspect, and mood.

³ Bybee's observed tendency is concerned with distance from the root and not explicitly about suffixing. (3) would also be accurate with PERSON, TAM, VOICE, and VALENCE placed directly to the left of ROOT in the case of Hupa.

e.g. valence and voice, were closer to the root than more productive morphemes that are less codependent on the root, e.g. TAM and subject marking. This was also observed at the word level, where morphemes found closest to the root in linear distance were also highest in mutual information to the root on average, with a sharp drop off with increasing distance, seen in Figure 1 (right).

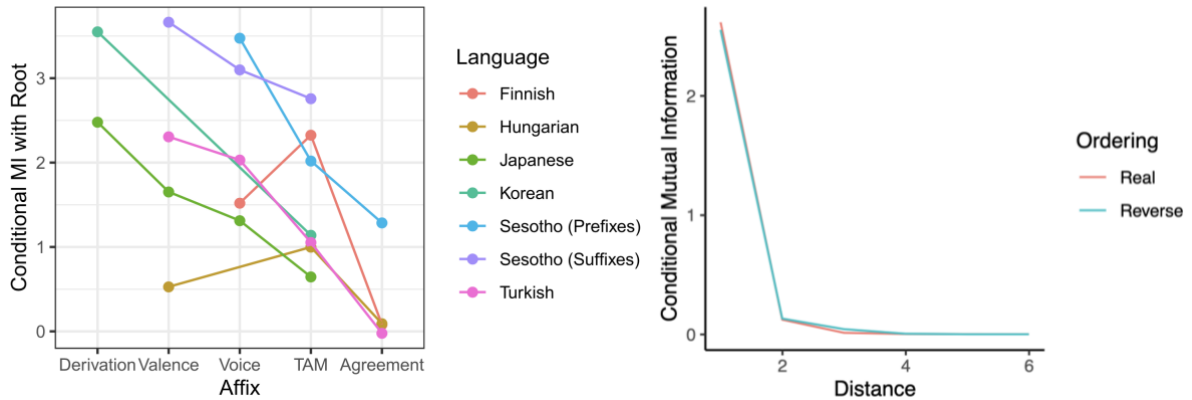


Figure 1. *Left*: Average root-conditional mutual information of morphemes grouped by position type (Hahn et al. 2021b). *Right*: Average root-conditional mutual information of Sesotho suffixes groups by their linear distance from the root (Hahn et al. 2021b)⁴

The motivation for this phenomenon was proposed by Hahn et al. (2021b) to be the memory-surprisal tradeoff, which states that placing mutually-informative morphemes closer together overall reduces the average surprisal the morpheme order has and thus maximizes information gain with smaller memory demands. As seen in Figure 2, Language A demonstrates a more efficient memory-surprisal tradeoff than Language B due to a morpheme order that is more aligned with information locality. As memory constraints increase on the x-axis, Language A has a morphological grammar that achieves a lower average surprisal earlier than Language B.

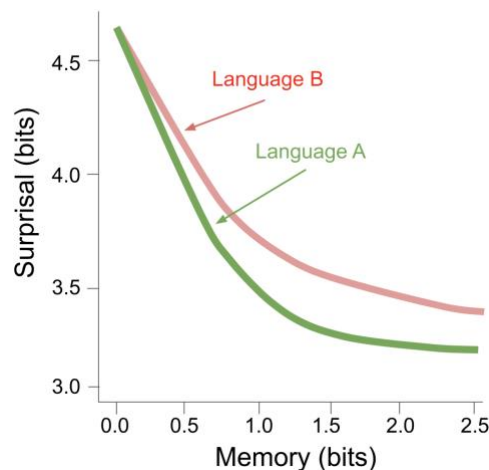


Figure 2. Example of memory-surprisal tradeoff curves of two hypothetical morphological grammars (Hahn et al. 2021b)

These results then raise the questions that are the main research topics of this project:

⁴ This figure is from the supplemental information of Hahn et al. (2021b).

1. How can these information-theoretic methods of Hahn et al. (2021b) be applied to a language like Hupa, given its low-resource nature and morphological intricacy?
2. What do the results of this application have in regards to its connection between Hupa and cross-linguistic morpheme tendencies and linguistic typology?
3. What applications can these results have on previous theoretical explanations and observations of Dene languages?

2. Data and methods.

2.1. CORPUS. All data for this project come from the Online Hupa Dictionary and Texts⁵. The data consist of a manual corpus of morphologically-segmented verb instances from Hupa narratives, all of which were told by Verdena Parker (Duval 2024).⁶ The corpus comprises of 2,827 verbs tokens extracted from 29 oral narratives. These verb tokens comprise of 239 unique verb root types and 113 unique prefix types. Table 2 provides a more detailed breakdown of the corpus, with token and type amounts for each morpheme template slot as well as the number of morpheme types that were omitted from slot-based measurements because their low frequency caused a null MI measurement. The data for this project was also not slotted, meaning there were not distinct columns for each slot and morphemes were instead segmented linearly.

	11 (Them)	11 (Adv)	10	9	8	7	6	5	4	3	2	1
Tokens	65	559	183	236	617	139	440	136	14	822	198	390
Types	3	14	1	2	5	3	13	3	1	7	7	3
Omitted	5	9	0	0	5	5	5	0	0	2	1	1

Table 2. Corpus token and type frequencies for each prefix position. The template slot numbers correlate to the numbering used in Table 1.

One nuanced aspect of this project is in the segmentation method of Hupa, since it does exhibit a fair amount of allomorphy and morphemic fusion. This project took a hybrid approach to this, where local, phonological variation and more transparent morphemic fusion were segmented as their underlying forms as seen in (5). (5a) is an example of the classifier *l* voicing to *l* after the first-person plural subject prefix *di*, which was segmented as its underlying, voiceless form. In (5b) the thematic and perfective prefixes *di* and *win*, respectively, fuse to form *de:* in the surface, but they are segmented as their underlying forms in this corpus. Homophonous morphemes were also recorded distinctly in the corpus.

- (5) a. *didilqos*
di-di-l-qos
 THEM-1.PL.SBJ-VAL-ROOT
 ‘We bite it off.’
- b. *a'de:ne'*
a:-'di-win-ne'
 ADV-3.ANIN.SBJ-THEM-PERF-ROOT
 ‘S/he said.’

As for stronger allomorphy, which often occurred with subject marking in perfective verb forms, these instances were recorded as their surface forms, as seen in (6a), where the first person singular subject prefix is recorded as its perfective form *e:* as opposed to its dictionary

⁵ <http://nalc.ucdavis.edu/dictionaries/hupa-lexicon.php>.

⁶ The data repository for this project can be found here: <https://osf.io/dkmzb/>.

form *wh*. In (6b) the first person plural subject prefix *di* reduces to *y* before the root *dil*. This was recorded in the corpus as its surface form. This segmentation approach was also the case with verb root variation, where aspectual and modal suppletion of the root was acknowledged in the corpus, but local allomorphy was not.

- | | | |
|-----|---|--|
| (6) | a. <i>se:la:n</i>
<i>si-e:-la:n</i>
PERF-1.SG.SBJ-ROOT
‘I was born.’ | b. <i>naydil</i>
<i>na:-y-dil</i>
ADV-1.PL.SBJ-ROOT
‘We go around.’ |
|-----|---|--|

2.2. SCRIPT AND DATA VISUALIZATION. Two primary scripts⁷ were utilized for this project. The first script measured the mutual information between morpheme pairs as well as the memory-surprisal tradeoff of the corpus. This script also had the ability to shuffle morpheme order to use as a baseline. The measurement of MI and surprisal from the script is based on the equation in (7) (Hahn et al., 2021b). The first fraction represents the overall morpheme sequence probability within the training section of the corpus by dividing the frequency of the morpheme sequence $F(w_0 \dots w_t)$ by the overall morpheme count in the training corpus $|C|$. The second fraction divides the conditional probability of the current morpheme w_t and its previous context $w_{0\dots t-1}$ by the conditional probability of the current morpheme and its previous context without the first morpheme w_0 in the held-out corpus, thus measuring the mutual information between the morpheme w_0 and w_t given the intervening morphemes $w_{1\dots t-1}$. This script measured the mutual information between all possible prefix types. However, this project is only concerned with prefixes’ mutual information to their root, hence only root-conditional mutual information was recorded.

$$(7) \quad I_t \equiv I[w_t:w_0|w_1 \dots w_{t-1}] \equiv \sum_{w_0\dots w_t} \frac{F(w_0\dots w_t)}{|C|} \log_2 \frac{P_t(w_t|w_0\dots w_{t-1})}{P_{t-1}(w_t|w_1\dots w_{t-1})}$$

Surprisal is measured from this equation by removing the denominator in the second fraction, which would then just give us the conditional probability of the morpheme w_t given its contextual sequence $w_{0\dots t-1}$, thus the surprisal. As for measuring memory constraints, (8) quantifies memory as the information content or mutual information (I_t) of T number of morphemes of each morpheme weighted by the morpheme’s distance in the sequence t (Hahn et al. 2021b).

$$(8) \quad M_T := \sum_{t=1}^T tI_t$$

A second script was used to optimize the morpheme order of the corpus. The algorithm first compiles all unique prefixes in the corpus and constructs a hypothetical word containing every prefix. Using the prefix’s MI measurements conditional to the root, it then iterates randomly through possible placements for each morpheme that would minimize surprisal. The following optimized forms were the result of running the algorithm for approximately 16 hours. Average MI and surprisal ratings were recorded then plotted in R using ggplot2 (Wickham 2016; R Core Team 2022).

3. Results.

⁷ Both scripts were graciously provided by Dr. Michael Hahn.

3.1. MUTUAL INFORMATION WITH LINEAR DISTANCE. Figure 3 visualizes the average root-conditional MI of morphemes grouped by their linear distance from the verb root. The real order appears to achieve a higher concentration of root-conditional MI closest to the root with a sharp drop off with increasing distance, a much sharper and consistent drop off than that of the randomized baseline.

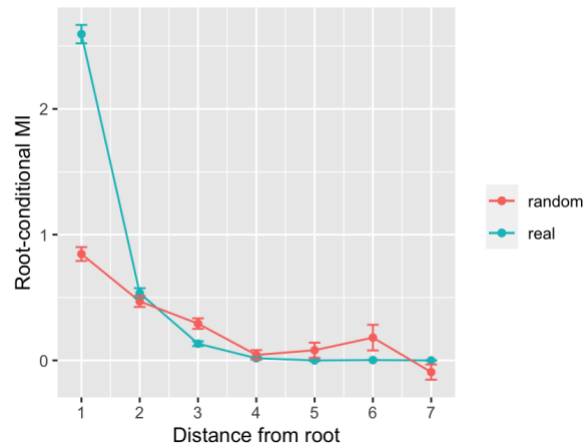


Figure 3. Average root-conditional MI of prefixes grouped by their linear distance from the root compared to randomized order baseline

3.2. MUTUAL INFORMATION WITHIN TEMPLATIC POSITIONS. MI measurements based on templatic position, seen in Figure 4, shed an interesting light that quantitatively confirms previous accounts of the complex Dene morphological template. We first see that the drop off in MI is not as consistent with distance as it is at the word level. Rather, certain morphemic placements seem to correlate with the unique aspects of Dene morphological order, such as aspectual and modal marking (A/M) being higher in root-conditional MI than first and second person marking (1/2) despite them being farther from the verb root. Also, thematic and adverbial markers are on average higher in root-conditional MI than their neighbors given their distance from the root.

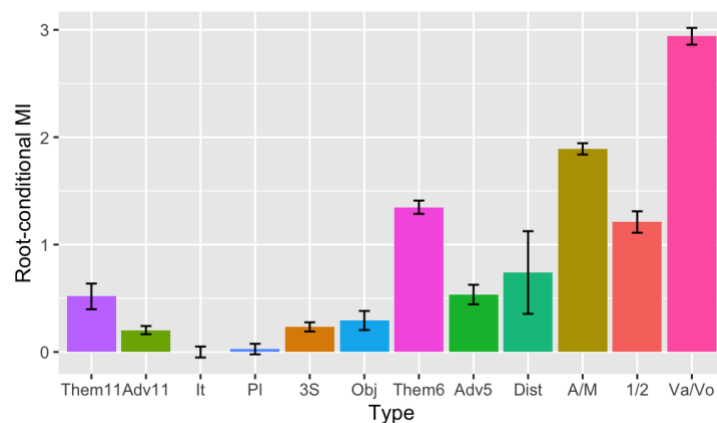


Figure 4. Average root-conditional MI of morphemes grouped by their templatic slot position⁸

⁸ Va/Vo = valence and voice (classifiers), 1/2 = first- and second-person subject, A/M = aspectual and modal, Dist = distributive, Adv5 = adverbial fifth slot, Them6 = thematic sixth position, Obj = object, 3S = third person subject, Pl = plural, It = iterative, Adv11 = adverbial 11th position, Them11 = thematic 11th position.

3.3. MEMORY-SURPRISAL TRADEOFF. Illustrated in Figure 5, the corpus’ real morpheme order achieved a lower average surprisal with increasing memory constraints than the randomized baseline. The optimized order expectedly achieved a lower surprisal than both the real and randomized orders. More importantly, the average surprisal achieved by the real morpheme order is more comparable to that of the optimized order than random.

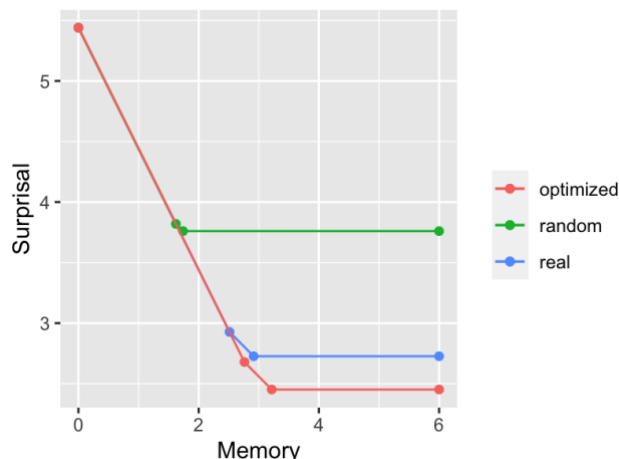


Figure 5. Memory-surprisal tradeoff curves for real, random, and optimized morphemes orders

4. Discussion.

4.1. ROOT-CONDITIONAL MI AND SURPRISAL. The emergent results from the project reveal a similar story to Hahn et al.’s research, specifically where word-level morpheme orders in Hupa correlated to ordering the most root-wise mutually-informative morphemes closest to the root. Also, this order achieved a lower average surprisal than randomized baselines and were more comparable to the surprisal achieved by optimized morphemes orders, a result that was also observed in the languages from Hahn et al (2021b), demonstrating a possible tendency of morpheme order to coincide more with efficiently-optimized orders than stochastic ones.

4.2. WORD-LEVEL VERSUS TEMPLATIC LEVEL. One of the more striking differences when comparing these results is that of word-level versus morpheme type mutual information. In other words, the proximal concentration of root-conditional MI with a stark that drop off that see at the word level, Figure 3, is not reflected when we look at the average MI of morphemes separated by their templatic position, Figure 4, while Hahn et al. observed a similar trend in MI drop off both at the word and position level. This could suggest the dichotomy between underlying and surface-level complexity, illustrating that while a language is theoretically quite complex in its verb template or hypothetical paradigms, real-world performance data and common verbs might still demonstrate comparatively-efficient structures to other languages.

4.3. OPTIMIZED ORDERS. The orders that the algorithm returned as optimized also reveals a fascinating correlation to Hupa and morphological typology. Figure 6 visualizes the modifications that the algorithm implemented on the different morpheme types on average. Because this project’s data were not slotted, the optimized template in Figure 6 was extrapolated by manually averaging the morpheme placements within each type in relation to the root. For example, if a morpheme slot contained three morphemes and the algorithm returned those three morpheme placements as a distance of one, four, and seven from the root, the optimized template will have that mean as its placement, hence the position four.

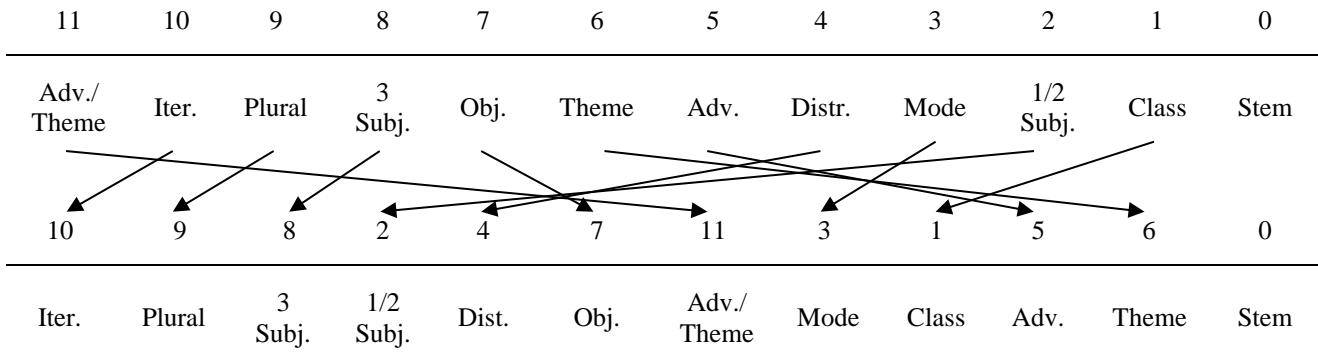


Figure 6. Real Hupa morpheme template (above) and the returned optimized orders calculated by the optimizing algorithm (below)

The most striking aspect of Figure 6 is the fact that the rearranged order that the algorithm returned now coincides more with previous cross-linguistic tendencies of morpheme order, specifically that previously observed by Bybee (1985a) and Greenberg (1963). In the optimized order, more derivationally-natured morphemes were moved closer to the verb root while more productive morphemes and those related to person marking were moved farther away from the root. This suggests that optimizing for processing efforts and minimizing surprisal can explain previously-observed morpheme ordering tendencies cross-linguistically.

4.4. REAL-WORLD AND THEORETICAL APPLICATIONS. The emerging results from this project reveal aspects of the statistical distribution of Hupa’s verbal morphology than can be applied to both formal and real-world research.

The statistical reality of morpheme order can partly describe the phenomenon of language acquisition of complex morphology. This has been observed in the nearly-accurate morpheme order in the speech of children acquiring Turkish and Mohawk, which were attributed to sensitivity to regularity patterns in the paradigm and knowledge of the lexicon, respectively (Aksu-Koç & Slobin 1985; Mithun 1989). These sensitivities to regular patterns or lexical items can be attributed to internalization of statistical relationships of morphemes to the root, hence their mutual information. This has been observed more specifically within the Dene language family in the acquisition of Navajo, with Courtney and Saville-Troiike (2002) observing that prefixes were acquired word-specifically and modal and thematic prefixes were either omitted or included root-dependently. Chee’s 2017 dissertation observed that verb roots and thematic markers were largely acquired first. However, the topic of language acquisition goes beyond statistical internalization and also involves perceptual salience and semantic relevance, which are proposed reasons as to why both Courtney and Saville-Troiike and Chee found that thematic markers farther from the root were acquired and mastered first before those closer to the root, which might not be immediately explainable using information theory alone. Mithun (1989) also attributes the order of syllable acquisition in Mohawk to both phonological prominence and importance to discourse.

An information-theoretic approach to morpheme order can also add interesting light to previous approaches. Bybee (1985a) attributes the observed tendency in (4) to be of semantic relevance to the root. However, she also acknowledges that the stronger semantic relevance a prefix has to the root, the lower its productivity usually is, which can be correlated to its co-occurrence to the root, hence its root-conditional mutual information. A more specific example is the order of first- and second-person versus third-person subject marking in Dene languages, a

discrepancy that Rice (2000) attributes to the different functions that these prefixes carry, namely syntactic or discourse-based. However, the stronger allomorphy of first- and second-person subject prefixes lend them to be more root-dependent in their surface forms, which could offer validity to their proximal location compared to third-person markers in this project's framework. This concept of larger allomorphy and phonological interaction closer to the root can extend to Campbell's (2007) bipartite model, which can also be related to how this allomorphy increases root-conditional MI of morphemes closer to the root. However, this project does not have an immediate answer to the question: Is this allomorphy driving proximity or is the proximity driving allomorphy?

Also, extensive discussion has been made on the ordering of subject and object agreement morphemes in languages which both Chomsky (1992) and Bybee (1985b) observe that object inflection is generally closer to the root. Chomsky attributes this to the syntactic proximity of objects to their verbs, which Bybee also claims but also possibly attributes it to object marking being semantically related to valence and voice. While this phenomenon can be partly explained in this project by the discrepancy in root-conditional MI between first-/second-person marking and object marking, there does not appear to be a significant MI difference between object and third-person subject marking. There is also not an immediate explanation as to why the algorithm returned an optimized order with first-/second-person marking farther from the root than object marking despite the former's far higher MI measurements.

5. Future directions and improvements. This project would benefit both from internal improvements to the data and methods as well as possible expansions to other sub-disciplines to linguistics.

As for the data, this project can always benefit from larger and more diversified data, especially given the number of prefix types that had to be omitted from the corpus given their minimal occurrences. Since this project only involved verb instances from narrative corpora, a possible improvement would involve adding more verb instances from different contexts, such as conversational data. Also, if this data were slotted, it would open more possibilities in the project regarding testing out different hypothetical morphological grammars as baselines and potentially improving morpheme order optimization.

Regarding future directions, it would be interesting to see this approach applied further to other low-resource, morphologically intricate, agglutinative languages to better assess how far-reaching information-theory can be used to describe cross-linguistic morpheme order, as well as possibly expand the existing work on information theory and non-concatenative morphology. Studies like these would also ideally be followed by incorporating psycholinguistic research to further analyze whether these computational models of statistical structure actually correlate to online processing of morphology. Lastly, the method with which one segments the morphemes in a word could have a great impact on the information-theoretic measurements of the corpus. This has been touched on in Hupa (Duval 2023), but further research should be done on the ideal segmentation methods for information theory and morphology and which method best reflects human cognition.

6. Conclusion. The original goal of this project was to examine to what extent information-theoretic concepts of mutual information and surprisal can apply to the analysis of morpheme order in the Dene language Hupa, since previous studies have only focused on higher-resource languages of Afro-Eurasia that already coincided with previously-observed morpheme order tendencies. The results of this project suggest that despite Hupa's morphological intricacy at the templatic level and smaller corpus size, word level morpheme orders in the corpus achieved

similar results to the languages analyzed by Hahn et al. in both mutual information to the root and average surprisal, which seem to have interesting implications to both Dene morphological theory and linguistic typology. This suggests that two languages could be comparatively efficient in their performance data despite the underlying, theoretical differences in their morphological grammars. Also, optimization results coincided with previously-noted morphological tendencies, which made the Hupa template more in line with typological observations of Bybee and Greenberg. Further innovations and diversifications of data would be beneficial to this project and to the expansion of these methods to other low-resource, agglutinative languages, as would the potential application of these results to real-world experiments of online morphological processing.

References

- Aksu-Koe, Ayhan A. & Dan I. Slobin. 1985. The acquisition of Turkish. In Dan I. Slobin (ed.), *The crosslinguistic study of language acquisition*, 839–878. New York: Psychology Press. <https://doi.org/10.4324/9781315802541-10>.
- Bybee, Joan L. 1985a. *Morphology: A study of relation between meaning and form* (Typological Studies in Language 9). Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.9>.
- Campbell, Amy. 2007. Hupa ditransitives and the syntactic status of R. Conference on Ditransitive Constructions. MPI-EVA, Leipzig.
- Chee, Melvatha R. 2017. *A longitudinal cross-sectional study on the acquisition of Navajo verbs in children aged 4 years 7 months through 11 years 2 months*. Albuquerque: The University of New Mexico dissertation.
- Church, Kenneth & Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16. 22–29. <https://aclanthology.org/J90-1003>.
- Courtney, Ellen H. & Muriel Saville-Troike. 2002. Learning to construct verbs in Navajo and Quechua. *Journal of Child Language* 29(3). 623–654.
- Duval, Cameron. 2023. Effects of different segmentation methods on information-theoretic morphology of Hupa verbs. Presented at Alberta Conference on Linguistics, Lethbridge. <https://doi.org/10.13140/RG.2.2.14921.76646>.
- Duval, Cameron. 2024. Information-theoretic applications to Hupa verbal morphology. OSF project. <https://doi.org/10.17605/OSF.IO/DKMZB>.
- Frank, Stefan, & Robin Thompson. 2012. Early effects of word surprisal on pupil size during reading. *Proceedings of the Annual Meeting of the Cognitive Science Society* 34. <https://escholarship.org/uc/item/411981tk>.
- Futrell, Richard, Peng Qian, Edward Gibson, Evelina Fedorenko & Idan Blank. 2019. Syntactic dependencies correspond to word pairs with high mutual information. In Kim Gerdes & Sylvain Kahane (eds.), *Proceedings of the Fifth International Conference on Dependency Linguistics*, 3–13. Paris: Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-7703>.
- Golla, Victor K. 1970. *Hupa grammar*. Berkeley: University of California dissertation. <https://escholarship.org/uc/item/6z0634f2>.
- Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Hale, John. 2001. A probabilistic Earley parser as a psycholinguistic model. *Second meeting of the North American Chapter of the Association for Computational Linguistics*. <https://doi.org/10.3115/1073336.1073357>.

- Hahn, Michael, Judith Degen & Richard Futrell. 2021a. Modeling word and morpheme order in natural language as an efficient trade-off of memory and surprisal. *Psychological Review* 128(4). 726. <https://doi.org/10.1037/rev0000269>.
- Hahn, Michael, Rebecca Mathew & Judith Degen. 2021b. Morpheme ordering across languages reflects optimization for processing efficiency. *Open Mind* 5. 208–232. https://doi.org/10.1162/opmi_a_00051.
- Hargus, Sharon. 1988. *The lexical phonology of Sekani*. London: Routledge. <https://doi.org/10.4324/9780429455285>.
- Hupa online dictionary and texts. 2008–2022. <http://nalc.ucdavis.edu/hupa/hupa-lexicon.php>.
- Kari, James. 1989. Affix positions and zones in the Athapaskan verb complex: Ahtna and Navajo. *International Journal of American Linguistics* 55(4). 424–454. <https://www.jstor.org/stable/1265241>.
- Kari, James. 1990. *Ahtna Athabaskan dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska.
- Kari, James. 1992. Some concepts in Ahtna Athabaskan word formation. In Mark Aronoff (ed.), *Morphology now*, 107–132. Albany, NY: SUNY Press.
- Kita, Kenji, Yasuhiko Kato, Takashi Omoto, & Yoneo Yano. 1994. A comparative study of automatic extraction of collocations from corpora: Mutual information vs. cost criteria. *Journal of Natural Language Processing* 1(1). 21–33. <https://doi.org/10.5715/jnlp.1.21>.
- Mithun, Marianne. 1989. The acquisition of polysynthesis. *Journal of Child Language* 16(2). 285–312. <https://doi.org/10.1017/S0305000900010424>.
- McDonough, Joyce. 2000. On the bipartite model of the Athabaskan verb. In Theodore Fernald and Paul Platero (eds.), *The Athabaskan languages: Perspectives on a Native American language family*, 139–166. Oxford: Oxford University Press.
- Monsalve, Irene F., Stefan Frank & Gabriella Vigliocco. 2012. Lexical surprisal as a general predictor of reading time. In Walter Daelemans (ed.), *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 398–408. Avignon: Association for Computational Linguistics. <https://aclanthology.org/E12-1041>.
- Pothos, Emmanuel M. & Patrick Juola. 2007. Characterizing linguistic structure with mutual information. *British Journal of Psychology*, 98(2). 291–304. <https://doi.org/10.1348/000712606X122760>.
- Potter, Brian. 1996. Minimalism & the mirror principle. *North East Linguistics Society* 26(1). 21. <https://scholarworks.umass.edu/nels/vol26/iss1/21>.
- R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Software package. Vienna. <https://www.R-project.org/>.
- Rice, Keren. 2000. *Morpheme order and semantic scope: Word formation in the Athapaskan verb*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511663659>.
- Smith, Nathaniel J. & Roger Levy. 2008. Optimal processing times in reading: A formal model and empirical investigation. *Proceedings of the Annual Meeting of the Cognitive Science Society* 30. <https://escholarship.org/uc/item/3mr8m3rf>.
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. <https://ggplot2.tidyverse.org>.