Tashlhiyt Intonation Alignment in MaxEnt

Aaron Kaplan
University of Utah

1 Introduction

One of the salient properties of MaxEnt grammars (Goldwater & Johnson, 2003) is that they assign a non-zero output probability to all candidates, even harmonically bounded ones. In other constraint-based frameworks such as Optimality Theory (OT; Prince & Smolensky, 1993/2004) and Harmonic Grammar (HG; Legendre et al., 1990), harmonically bounded candidates are completely inaccessible: they are never possible outputs under any ranking/weighting, even in probabilistic versions of these theories such as Stochastic OT (Boersma, 1997), Partial Orders (Anttila, 1997), and Noisy HG (Boersma & Pater, 2016). MaxEnt therefore offers a unique opportunity to probe formal differences between simple and collective harmonic bounding, which are illustrated in (1 a) and (1 b), respectively.

(1) (a)

	C_1 x	C_2 y	H
a. BOUNDER		-1	$-y$
b. BOUNDED	-1	-1	$-x-y$

(b)

	C_1 x	C_2 y	H
a. BOUNDER	-2		$-2x$
b. BOUNDED	-1	-1	$-x-y$
c. BOUNDER		-2	$-2y$

In these tableaux, x and y represent constraint weights. Simple harmonic bounding occurs in (1 a), where BOUNDER has a proper subset of BOUNDED's violations, and BOUNDER cannot win under any ranking or weighting, assuming only positive weights.[1] A weight of zero for C_1 leads to a tie, but BOUNDED can never outperform BOUNDER. See Keller (2000), Boersma & Pater (2016), and Hayes & Kaplan (to appear) for discussion of ties and zero-weighted constraints.

In (1 b), BOUNDED is collectively harmonically bounded (Samek-Lodovici & Prince, 1999) by the other two candidates. Neither BOUNDER has a proper subset of BOUNDED's violations, but BOUNDED cannot outperform both candidates at once. If $x > y$, BOUNDED has a better score than candidate (a) but a worse score than candidate (c). If $x < y$, it has a better score than (c) but a worse score than (a). If $x = y$, the three candidates have identical scores.

* I am grateful to AMP 2024 participants for their feedback on this work. I am especially grateful to Bruce Hayes for extensive discussion of issues related to the topic of this paper and to Timo Roettger for generously sharing his Tashlhiyt data with me.

[1] Negative weights can subvert harmonic bounding (Pater, 2009), and much work in HG therefore assumes that weights must not be negative. I make the same assumption here.

Proceedings of *AMP 2023/2024*

Let us assume that OT and HG are correct to distinguish harmonically bounded candidates from other candidates by rendering the former impossible outputs while permitting the latter to win given an appropriate ranking/weighting. MaxEnt offers a way to approximate this situation in (1 a) but not (1 b). In (1 a), the probability of BOUNDED approaches zero as x increases. While this probability never actually reaches zero, it can be placed arbitrarily close to that value so that for practical purposes, BOUNDED has no chance of being selected as an output. But in (1 b), because BOUNDED's score must be between those of the other candidates, its probability must be, too. As its probability approaches zero (which can be contrived by increasing the weight of either constraint), so does the probability of one BOUNDER. It is impossible to (nearly) exclude BOUNDED without doing the same to a BOUNDER.

What is the source of the different behavior of these tableaux under MaxEnt? Two possibilities present themselves. First, this could be an artifact of simple versus collective harmonic bounding. But there is another potentially relevant difference between the tableaux: only in (1 a) does BOUNDED have a violation not shared by any other candidate. Perhaps this property, independent of harmonic bounding, is what allows MaxEnt to effectively eliminate one candidate while leaving others' viability intact.

Intonational alignment in Tashlhiyt Berber allows us to test these competing explanations in the context of actual linguistic data. On the basis of an analysis of Tashlhiyt intonation developed below, I will argue that there is indeed a fundamental difference between simple and collective harmonic bounding. MaxEnt can exclude simply harmonically bounded candidates even in the absence of unshared violations.

2 Contour Alignment and Dislocation in Tashlhiyt Berber

Grice et al. (2015) and Roettger (2017) investigate the phonetic and phonological properties of intonational contours that mark polar questions (2) and contrastive statements (3) in Tashlhiyt Berber (Afro-Asiatic; Morocco). They report that these contours preferentially fall on the final syllable of an utterance, as indicated by underlining in the examples below.

(2) (a) is i-nna tiri
INT 3MASC.SG-say she wanted
'Did he say 'she wanted'?'

(b) is i-nna tn̩zam
INT 3MASC.SG-say you were sold
'Did he say 'you were sold'?'

(c) is i-nna tugl̩
INT 3MASC.SG-say she hanged
'Did he say 'she hanged'?'

(3) (a) i-nna tilit
3MASC.SG-say she has it
'He said 'she has it'.'

(b) i-nna tm̩di
3MASC.SG-say she trapped
'He said 'she trapped'.'

(c) i-nna tudˁn̩
3MASC.SG-say she suffered
"He said 'she suffered'.'

The intonational contour can be dislocated from its final position. Perhaps the most dramatic example of this (and the context in which harmonic bounding will arise) occurs when the phrase-final word consists solely of voiceless obstruents, as in (4).

(4) is i-nna tk̩ʃf̩
INT 3MASC.SG-say it dried
'Did he say 'it dried'?'

Words such as /tkʃf/ ‘it dried’ present a conundrum because without voicing, they can host no intonational content. Grice et al. (2015) report that in constructions like (4), speakers exhibit the three strategies exemplified in (5). An epenthetic vowel may appear in the final syllable to host the contour (5 a), the contour may be shifted to the preceding word (5 b), or the contour may be unrealized altogether (5 c). Grice et al. report no frequency data for these particular patterns, but they state that there is intraspeaker variation: one speaker may use all three strategies.

(5) (a) is inna tk̩ʃəf
    (b) is inna tk̩ʃf̩
    (c) is inna tk̩ʃf̩

Even when the phrase-final word is a more hospitable host for the intonational contour, displacement from the final syllable to the penultimate syllable is still possible. Grice et al. identify three factors that affect the likelihood of this displacement. The first is syntactic context: the contour is more likely to appear on the final syllable in polar questions (2) than in contrastive statements (3).

The second factor is sonority: the contour is attracted to the syllable with the most sonorous nucleus. In (3 a), both the final syllable and the penultimate syllable contain a vowel. In (3 b), the final syllable’s nucleus is a vowel and the penultimate syllable’s nucleus is a sonorant consonant. All four of Grice et al.’s (2015) speakers produced a final-syllable contour in the latter context more often than in the former. In (3 c), the penult’s nucleus is a vowel and the final syllable’s nucleus is a sonorant consonant. Three of the four speakers produced a final-syllable contour less often in that context than in the vowel/vowel context of (3 a); see section 3 for discussion. Similar patterns hold for (2), though ceiling effects arise because of polar questions’ greater tendency to exhibit final-syllable contours. Grice et al. report that of the configurations shown in (2), only (2 c) exhibited variation in the contour’s placement. When a polar question’s final syllable contained a vowel, as in (2 a) and (2 b), their speakers invariably placed the contour on that syllable.

Finally, heavy syllables also attract the contour. For polar questions, all four of Grice et al.’s speakers invariably produced a final-syllable contour when that syllable was heavy. For contrastive statements, three of the four speakers produced a final-syllable contour more often when that syllable was heavy than when it was light.

Grice et al. do not disaggregate the effects of sonority and syllable weight, so it is impossible to compare, for example, the rate of final contour placement in (3 a) to that of (6), which is identical except that the final syllable is light.

(6)	i-nna	tili
	3MASC.SG-say	she has
	‘He said ‘she has’.’	

For this reason, the analysis developed in the next section accounts for only the effects of syntactic context and sonority.

3 Analysis

The analysis begins with the data in (5) because that is where the relevant harmonic bounding relationships will emerge. Three constraints are needed. Two are straightforward: DEP penalizes the vowel epenthesis seen in (5 a), and MAX-Contour penalizes candidates that do not realize the intonational contour (5 c). A third constraint, defined in (7), favors placement of the contour as close to the right edge of the utterance as possible.

(7) ALIGN-R: Assign one violation for each syllable between the intonational contour and the right edge of the utterance.

The constraints in (8) are not necessary to account for (5), but they anticipate other data to be included in the analysis. COINCIDE(Contour, V) captures the sonority effect described above by encouraging the intonational contour to appear on a vowel. (See Zoll (1998) for discussion of COINCIDE constraints.) ALIGN-R-PQ adds an extra incentive for final-syllable placement of the contour in polar questions. It operates exactly as ALIGN-R does, but it assigns violations only in the context of polar questions.

(8) (a) COINCIDE(Contour, V): Assign one violation for each intonational contour that does not surface on a syllable with a vocalic nucleus.

(b) ALIGN-R-PQ: In a polar question, assign one violation for each syllable between the intonational contour and the right edge of the utterance.

In the tableaux below, attested forms are marked with (☞). Weights and probabilities were derived with OTSoft (Hayes et al., 2013). Because frequencies are unavailable for (5), the three licit outputs were all assigned identical frequencies in the OTSoft input file. That input file and other files for different conditions presented below are available here: https://github.com/afkaplan/Tashlhiyt-tonal-alignment.

(9)

/is inna tkʃf/	MAX-Cont 20.083	DEP 20.083	ALIGN-R 9.196	COINCIDE (Cont,V) 9.152	ALIGN -R-PQ 0.842	H	p
(☞) a. is inna tḳʃ<u>əf</u>		−1				−20.083	.333
(☞) b. is in<u>na</u> tḳʃf̣			−2		−2	−20.077	.335
(☞) c. is inna tḳʃf̣	−1					−20.083	.333
d. is inna tḳʃəf	−1	−1				−40.167	.000
e. is in<u>na</u> tḳʃəf		−1	−2		−2	−40.160	.000
f. is inna <u>tək</u>ʃf̣		−1	−1		−1	−30.122	.000

This tableau combines properties of (1 a) and (1 b). Candidates (d)–(f) are simply harmonically bounded just like BOUNDER from (1 a). But this time, none of those candidates has a violation that is not shared by a licit candidate, and in this way the tableau resembles (1 b). For example, candidate (d) has exactly the union of the violations of candidates (a) and (c), and it is therefore simply harmonically bounded by each of them while having no violation that is not shared by one of them. Candidate (e) stands in the same relationship with candidates (a) and (b). Candidate (f) is simply harmonically bounded by candidate (a), and its violations of ALIGN-R and ALIGN-R-PQ are shared by candidate (b).

Even though the harmonically bounded candidates lack unshared violations, MaxEnt assigns only the licit candidates non-negligible probabilities as indicated in the column labeled p. Apparently, simple harmonic bounding provides the tools for this outcome on its own. The reason for this seems to be that candidates (d)–(f) violate two of the following constraints each while candidates (a)–(c) violate one each: MAX-Contour, DEP, and ALIGN-R. (COINCIDE(Contour, V) is inactive in this tableau, and ALIGN-R-PQ merely reinforces the penalties assigned by ALIGN-R.) Consequently, increasing the weights of these three constraints increases the penalty incurred by all six candidates, but candidates (d)–(f) are doubly harmed. MaxEnt simply needs to raise those weights enough to render the illicit candidates sufficiently improbable. As long as MAX-Contour and DEP have equal weights, and as long as their weights are twice the summed weights of ALIGN-R and ALIGN-R-PQ (to account for candidate (b)'s two violations of the ALIGN constraints), the probability space will be equally divided among the three licit candidates.

This strategy for dealing with illicit harmonically bounded candidates can be embedded in a larger account of intonational alignment, as (10) shows. In these tableaux, the "Target" column contains the frequencies reported by Grice et al. (2015:250) for the combined data from all four of their speakers. Individual speaker behavior is considered below. The form in (10 a) is a contrastive statement, so ALIGN-R-PQ is inactive. This form is from (3 c), and it instantiates the effect of sonority on contour placement because the penultimate syllable's nucleus is more sonorous than the final syllable's nucleus. Consequently, COINCIDE(Contour, V) and ALIGN-R conflict in their adjudication of the two licit outputs, and when the former has a slightly lower weight than the latter, candidate (a) is slightly more probable than candidate (b). The tableau in (10 b) is just like (10 a) except that it is a polar question (2 c). ALIGN-R-PQ provides more incentive for a final-syllable contour in this case. In both tableaux, predicted frequencies match the target frequencies. Taken as a whole, then, the weights used in (9) and (10) model the data reported by Grice et al. and assign illicit candidates—including those that are harmonically bounded—probabilities of approximately 0.

(10) (a)

/inna tudˤn/	MAX-Cont 20.083	DEP 20.083	ALIGN-R 9.196	COIN-CIDE (Cont,V) 9.152	ALIGN-R-PQ 0.842	*H*	*p*	Target
(☞) a. inna tudˤn̩				−1		−9.152	.511	.511
(☞) b. inna tudˤn̩			−1			−9.196	.489	.489
c. inna tudˤn̩			−2			−18.393	.000	.000
d. inna tudˤn̩	−1					−20.083	.000	.000
e. inna tudˤən		−1				−20.083	.000	.000
f. inna tudˤən		−1	−1			−29.280	.000	.000

(b)

/inna tugl/	MAX-Cont 20.083	DEP 20.083	ALIGN-R 9.196	COINCIDE (Cont,V) 9.152	ALIGN-R-PQ 0.842	*H*	*p*	Target
(☞) a. inna tugl̩				−1		−9.152	.708	.708
(☞) b. inna tugl̩			−1		−1	−10.039	.292	.292
c. inna tugl̩			−2		−2	−20.077	.000	.000
d. inna tugl̩	−1					−20.083	.000	.000
e. inna tugəl		−1				−20.083	.000	.000
f. inna tugəl		−1	−1		−1	−30.122	.000	.000

The target frequencies in these tableaux combine data from the productions of four people and may therefore not reflect the grammar of any actual speaker of Tashlhiyt. I submitted the data from each of the four speakers to OTSoft. The results are below. Table 1 shows the weights assigned to each constraint for each speaker, and Table 2 shows how well a MaxEnt grammar with these weights matches the frequency data reported by Grice et al. for constructions like (2 c) and (3 c). For readability, values that round to 0.000 are replaced with '—' in Table 2. Recall that there is no frequency data for voiceless obstruent-only target words such as /tkʃf/. The fact that MaxEnt consistently assigns all and only the attested forms for that item a probability substantially different from zero indicates success here; the 0.333 target is itself meaningless.

For the two items for which frequency data is available, MaxEnt often precisely replicates those patterns. Where it does not do so, that is because the speaker's behavior deviates from the general trends that the constraints are designed to capture. S3 produces more final-syllable contours in contrastive statements than in polar questions; the overall trend in the opposite direction motivated ALIGN-R-PQ. Not surprisingly, MaxEnt assigns ALIGN-R-PQ a weight of zero in this case, effectively turning it off and making it impossible to distinguish /inna tudˤn/ from /inna tugl/ analytically. The analysis therefore predicts identical patterns for those two forms and yields probabilities that are the average of the two forms' target probabilities.[2]

S2 never produces a final-syllable contour for /inna tudˤn/, and consequently COINCIDE(Contour, V) outweighs ALIGN-R for this speaker alone. As a result, *[inna tudˤn̩] is assigned a non-negligible probability because its contour appears on a vowel, and [inna tudˤn̩] (a licit option in general, though not one produced by this speaker) also has a non-negligible probability as a byproduct of ALIGN-R being weighted high enough to give other candidates with final-syllable contours a fighting chance. Perhaps this speaker provides evidence for separate COINCIDE(Contour, V) constraints, one for contrastive statements and another for polar questions. For some speakers the two constraints are weighted similarly, but for others, like this one, they have very different weights.

[2] When an ALIGN-R constraint specific to contrastive statements is added to the analysis, MaxEnt matches the target frequencies perfectly.

	MAX-Cont	DEP	ALIGN-R	COINCIDE(Cont, V)	ALIGN-R-PQ
S1	19.731	19.731	8.995	8.659	0.868
S2	22.246	22.246	5.279	10.564	5.844
S3	20.100	20.100	10.045	9.707	0.000
S4	18.858	18.858	8.123	7.024	1.304

Table 1: MaxEnt weights for each speaker

		S1		S2		S3		S4	
		p	Target	*p*	Target	*p*	Target	*p*	Target
/is inna tkʃf/	is inna tk̩ʃəf	0.333	0.333	0.333	0.333	0.332	0.333	0.333	0.333
	is inna tk̩ʃf̩	0.333	0.333	0.334	0.333	0.336	0.333	0.334	0.333
	is inna tk̩ʃf̩	0.335	0.333	0.333	0.333	0.332	0.333	0.333	0.333
	is inna tk̩ʃəf	—	—	—	—	—	—	—	—
	is inna tk̩ʃəf	—	—	—	—	—	—	—	—
	is inna təkʃf̩	—	—	—	—	—	—	—	—
/inna tudˤn/	inna tudˤn̩	0.583	0.583	0.005	—	0.584	0.667	0.750	0.750
	inna tudˤn̩	0.417	0.417	0.990	1.000	0.416	0.333	0.250	0.250
	inna tudˤn̩	—	—	0.005	—	—	—	—	—
	inna tudˤn̩	—	—	—	—	—	—	—	—
	inna tudˤən	—	—	—	—	—	—	—	—
	inna tudˤən	—	—	—	—	—	—	—	—
/inna tugl/	inna tugl̩	0.769	0.769	0.636	0.636	0.584	0.500	0.917	0.917
	inna tugl̩	0.231	0.231	0.364	0.364	0.416	0.500	0.083	0.083
	inna tugl̩	—	—	—	—	—	—	—	—
	inna tugl̩	—	—	—	—	—	—	—	—
	inna tugəl	—	—	—	—	—	—	—	—
	inna tugəl	—	—	—	—	—	—	—	—

Table 2: Predicted and target frequencies for each speaker using the MaxEnt weights in Table 1. '—' indicates a value that rounds to 0.000.

This analysis captures only part of the grammar of intonation alignment in Tashlhiyt. Further development requires data that is not presently available concerning the frequencies of the variants in (5) and the separate effects of syllable weight and sonority.

To return to the issue of harmonic bounding, MaxEnt consistently assigns a probability near zero to the simply harmonically bounded candidates at issue in (9). In at least some cases, then, MaxEnt can mimic the OT/HG treatment of simple harmonic bounding, even when the bounded candidates' violations are all shared by licit outputs. Though the analysis is far from a comprehensive treatment of Tashlhiyt, the results presented here suggest that this exclusion of harmonically bounded candidates is durable enough to persist in a larger and more complex analysis.

4 Conclusion

Assuming it is desirable to exclude harmonically bounded forms from the output set, collective harmonic bounding presents a challenge to MaxEnt that simple harmonic bounding does not. In the cases examined here, simply harmonically bounded candidates can be (nearly) excluded simply because they have more violations than (some) other candidates. Collective harmonic bounding does not present that situation and in fact partly reverses it because each bounder has more violations of a relevant constraint than the bounded candidate. Simple and collective harmonic bounding introduce different violation asymmetries, and this has consequences for the probability relationships between candidates that they allow in MaxEnt.

There are, of course, ways to effectively eliminate collectively bounded candidates in MaxEnt without taking a bounder with them. If we add to (1 b) a constraint that penalizes only BOUNDED, we provide MaxEnt a tool to decrease that candidate's probability to as close to zero as we like. This is illustrated in (11): BOUNDED remains collectively harmonically bounded, and its probability decreases as z increases. In that sense C_3 in this tableau plays the same role that C_1 plays in (1 a).

(11)

	C_1 x	C_2 y	C_3 z	H
a. BOUNDER	-2			$-2x$
b. BOUNDED	-1	-1	-1	$-x-y-z$
c. BOUNDER		-2		$-2y$

But C_3's consequences for BOUNDED are independent of harmonic bounding. Any candidate's probability can be manipulated in this way if it is the sole candidate that violates a particular constraint—that is, if it has an unshared violation (or at least a particular kind of unshared violation). What makes C_1 from (1 a) unique—and ambiguous—is that the violation it assigns to BOUNDED is both unshared and the reason BOUNDED is simply harmonically bounded. What the analysis of Tashlhiyt developed here shows is that the latter property is sufficient: simply harmonically bounded candidates can be effectively ruled out in MaxEnt even in the absence of unshared violations. Furthermore, that can be accomplished in the context of an analysis of a wider set of facts.

References

Anttila, Arto (1997). Deriving variation from grammar. Hinskens, Frans, Roeland van Hout & W. Leo Wetzels (eds.), *Variation, Change and Phonological Theory*, John Benjamins, Philadelphia, vol. 146 of *Current Issues in Linguistic Theory*, 35–68.

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences*, University of Amsterdam, Amsterdam, vol. 21, 43–58.

Boersma, Paul & Joe Pater (2016). Convergence properties of a Gradual Learning Algorithm for Harmonic Grammar. McCarthy, John J. & Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, Equinox, Bristol, CT, 389–434.

Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Workshop on Variation within Optimality Theory*, Stockholm University, 113–122.

Grice, Martine, Rachid Ridouane & Timo B. Roettger (2015). Tonal association in Tashlhiyt Berber: Evidence from polar questions and contrastive statements. *Phonology* 32, 241–266.

Hayes, Bruce & Aaron Kaplan (to appear). Zero-weighted constraints in noisy Harmonic Grammar. *Linguistic Inquiry* .

Hayes, Bruce, Bruce Tesar & Kie Zuraw (2013). OTSoft 2.6. software package, URL `http://www.linguistics.ucla.edu/people/hayes/otsoft/`.

Keller, Frank (2000). *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Ph.D. thesis, The University of Edinburgh.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky (1990). Harmonic Grammar – a formal multi-level connectionist theory of linguistic well-formedness: An application. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum, Cambridge, MA, 884–891.

Pater, Joe (2009). Weighted constraints in generative linguistics. *Cognitive Science* 33:6, 999–1035.

Prince, Alan & Paul Smolensky (1993/2004). Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder. Published 2004, Malden, MA: Blackwell.

Roettger, Timo B. (2017). *Tonal Placement in Tashlhiyt: How an Intonation System Accommodates to Adverse Phonological Environments*. Studies in Laboratory Phonology, Language Science Press, Berlin.

Samek-Lodovici, Vieri & Alan Prince (1999). Optima. ROA-363, Rutgers Optimality Archive, `http://roa.rutgers.edu`.

Zoll, Cheryl (1998). *Parsing below the Segment in a Constraint-Based Framework*. CLSI Publications, Stanford, CA.