A Case of OCP Effects in Intonational Phonology

Byron Ahn (byron-[æt]-ucla-[dat]-edu)
UCLA Department of Linguistics

Abstract

Assuming a set of tone-meaning mappings for English intonational phonology such as Pierrehumbert and Hirschberg (1990), certain data contain tones that would be otherwise unexpected. That is to say, we find surface representations that seem to contradict the assumption that the choice of tone target in English intonation is meaningful.

Given such data, one might assume that a theory of tone-meaning mappings is insufficient. However, I will attempt to show that, by modeling intonational phonology with a system of Optimality Theoretic constraint, we can derive these seemingly aberrant surface representations as simple cases of OCP effects.

1. Introduction

1.1. ToBI Model of Intonation

Throughout this paper, I will assume the Tones and Break Index (ToBI) model of Mainstream American English (MAE_ToBI), as originally proposed in Silverman et al. 1992, and most recently updated in Beckman and Ayers (1997). The MAE_ToBI model is comprised of five pitch accents, two phrase accents, and three boundary tones, which are laid out in (1).

(1) *Tone Inventory of MAE_ToBI*

Pitch Accents[1]: L*, H*, L+H*, L*+H, H+$^{!}$H*
Phrase Accents: L-, H-
Boundary Tones: L%, H%, %H

The diacritics (*, -, %) are not themselves a part of the tone inventory. Instead they are markers of different phonological objects; L* is not the same kind of phonological object as L-, which is not the same as L%. We can imagine that T*, T- and T% are three inherently different classes of tone targets in the same way that vowels and consonants are two different classes of segments.[2]

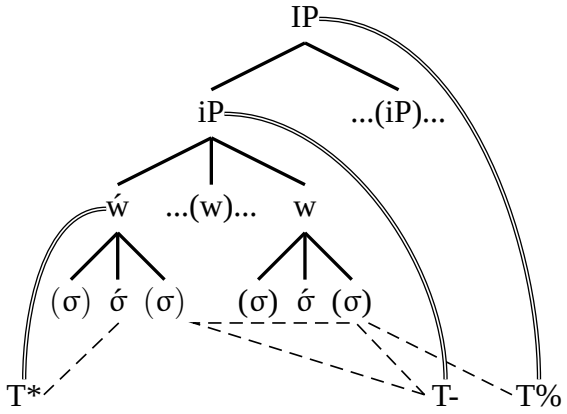
The reason for positing that there are different tone targets types is that each one can be targeted for linguistic generalizations, such as the following. Pitch accents are linked to prominent words (PromWd) and are realized on the stressed syllable(s). Phrase accents are linked to the right edge of an intermediate phrase, and are realized on the final syllable of that intermediate phrase (ip), spreading to the right edge of the last PromWd. Boundary tones are linked to the edge of an

[1] L* and H* are monotones, and the rest are bitones. It should be noted that there is no direct correlation between bitones and contour tones in the sense that the target syllable of the bitone does not necessarily have a contour within it – instead, one target of the bitone may be on a neighboring syllable.

[2] If we were to take an autosegmental approach, we could say that T*s, T-s and T%s each have their own tier, in the same way that vowels and consonants are said to each have their own tier.

intonation phrase (IP), and are realized on only the syllable aligned with that IP edge. All of these facts are summarized in the diagram in (2). Note that ẃ is shorthand for prominent word, and that the dotted lines represent the domain to which the tone is linked.

(2) *Model of MAE_ToBI*



It should be noted that there is at least one syllable in this diagram (the second unstressed syllable of the PromWd) is not tonally specified. When such a case occurs, the tonal value of the underspecified material is determined by linearly connecting tone targets, widely called interpolation. That is, in the example in (2), interpolation takes place, connecting the tone targets of the T* and the left edge of the T- domain.

If the first tone target is not the first syllable, there will be underspecified syllables without a target on the left to interpolate with. This is resolved by the fact that a medium-range target is assigned to the first syllable, from which interpolation to the next tone target can take place.

1.2. Tone and Meaning

English intonation is striking with regard to its large inventory of tones. How, for each time a T is required by the grammar, does the grammar choose the proper tone? It has been claimed that each tonal element of the grammar is associated with specific interpretations. For example, “the items made salient by the H* are to be treated as ‘new’ in the discourse.” Pierrehumbert and Hirschberg (1990)

Given any such theory which maps tone to meaning, it must be the case that the tone is underlyingly assigned to units that bear the relevant meaning. To be more explicit, tones must be targeting meaningful units such as phrases, words, or morphemes, but not phonological units such as foot or syllable. This fact is reflected in Gussenhoven (2004) which makes the broader point that prosodic constituents are not determined just phonologically, but with reference to syntax, semantics and information structure.

That said, given that tones must be realized on phonological units, we require a grammar to map the underlying representations to phonetic realizations. I will formulate such a grammar in terms of OT constraints which have the goals laid out in (3).

(3) *Work to be Done in the Grammar*[3]

a. Define and place domain of prominence, given a focus domain
b. Place T* in a PromWd, in a one-to-one ration to its stressed syllable(s)

c. Place T- at the end of a ip, and spread it appropriately

It should be noted that the constraints described in (3) will not be making too many changes from the underlying form to the surface intonational representation. That is, an optimal candidate often has no constraint violations. According to Gussenhoven (2004), this is true for most systems of intonation – though there are exceptions (e.g. Bengali OCP, Hayes and Lahiri 1991).

2. Intonational OT Grammar

In this section, I formulate the grammar that will accomplish the goals set forth in (3). However, due to the fact that the optimal candidate usually (perhaps, always) violates no constraints, ranking the constraints nearly impossible. We will therefore not rank any constraints or have any tableaux until a ranking can be made; until then, we can just as well assume that these constraints are inviolable.

2.1. Prominent Domain

As mentioned earlier, prominent words are crucially important for determining pitch accent placement. However, how is the unit of the PromWd itself defined? First of all, PromWds are crucially the size of a prosodic word (ProsWd), even when more than just a word is being focused.[4] This is, in part, due to the constraint in (4).

(4) PROMWD=PROSWD
A PromWd is equal in size to a prosodic word. Assign a violation each time it is not.

As we said earlier, prosodic structure can be determined by the information structure. A common reference to information structure is that of a focus domain. In particular, we can use focus domains to predict which words are a PromWd. Notice that in (5)–(7), different focus domains yield the same prominent word.[5]

(5) A: What happened last night?
B: [They watched PSYCHO]$_{foc}$.

(6) A: What did they do last night?
B: They [watched PSYCHO]$_{foc}$.

(7) A: What did you watch last night?
B: They watched [PSYCHO]$_{foc}$.

[3]This grammar must be expanded in many ways; however, these constraints cover the phenomena presented in this paper. For a discussion of further necessary constraints, see Further Research.

[4]There is the case of marking a morpheme as prominent, as in "She wrote it and then she **RE**wrote it." Here it seems that the PromWd is smaller than a ProsWd. Presumably, this could be derived through correctly ranking constraints.

[5]They may yield the same prominent word, though when focusing an XP, it is also common for X to be focused as well as the complement of X. For example, the VP in (6) is focused, and we see the complement of V as prominent; but an alternative pattern might be "They [WATCHED PSYCHO]$_{foc}$." However, it would be infelicitous to only make the V prominent when the entire VP is focused, as in "#They [WATCHED psycho.]$_{foc}$". For a more full discussion of this, see Selkirk (1996), Welby (2003).

The generalization arising from data like (5)–(7) is that focus domains have a PromWd at their right edge. Moreover, the responses in (5)–(7) would be ungrammatical if there were no PromWd at the right edge. I formalize this in (8).

(8) ALIGN(FOC-CONSTITUENT, R; PROMWD, R)[6]
For every focused constituent, align a PromWd to its right edge. If there is no PromWd aligned with the right edge of a focus constituent, assign a violation for every syllable between its right edge and the right edge of a PromWd.

2.2. Place T*s in a PromWd

Pitch accents are realized on the stressed syllable(s) of a PromWd. I term the relevant phonological bearer of pitch accents T* Bearing Unit (T*BU), whose size is regulated by the constraint in (9).

(9) T*BU=PROMWDSTRESSSYLL
*Every stressed syllable of a PromWd is a pitch accent bearing unit (T*BU). Assign a violation if T*BU is not a stress syll.*

I assume that pitch accent choice is determined by another set of constraints, which make explicit reference to syntax, semantics and information structure. Furthermore, I assume that these constraints mark PromWds with the specific T*. The T* marking the PromWd is mapped to its T*BUs with the constraints laid out in (10)–(13). Moreover, optimally, there is a one-to-one-to-one relationship between T*s, association lines, and T*BUs. This generalization can be decomposed in to two major sections. First, for every T*BU optimally has a single T associated with it.

(10) ASSOC(T*BU, T*)
*For every T*BU, ensure an association line links it to a T* of the kind that underlyingly marks the PromWd. Assign a violation if T*BU does not have a T* associated with it.*

(11) NOCROWD(T*BU)
*Only one T of any kind – pitch accent, phrase accent or boundary tone – per T*BU. Assign a violation if more than one T is realized on a T*BU.*

Second, every T* is optimally associated with one T*BU.

(12) ASSOC(T*, T*BU)
For every T, ensure an association line links it with a T*BU. Assign a violation if T* is associated with anything but T*BU, or if it is not associated to anything.*

(13) NOSPREAD(T*)
The domain of T can be no larger than T*BU. Assign a violation if T* is realized on more than just a T*BU.*

[6]Note that PromWds may also exist outside of a focus domain; their placement is not addressed in this paper.

2.3. Place and Spread T-

For T-, similar to T*, I assume that constraints sensitive meaningful or structural contexts militate the choice of H- versus L-. Again similar to T*, I assume that these constraints mark the ip with this T-, and my constraints yield the domain of realization of the T-.

(14) ASSOC(T-,]ip)
For every T-, associate it to the right edge of the ip. Assign a violation for every syllable between the associated element and the ip's right edge.

(15) ALIGN(T-, R; ip, R)
For every T-, align its right edge with the ip's right edge. Assign a violation for every syllable between the the right edge of T- and the ip's right edge.

(16) ALIGN(T-, L; PROMWD, R)
T- spreads leftward to the right edge of the last PromWd. Assign a violation for every syllable between the the left edge of T- and the PromWd's right edge.

If there are multiple PromWds, we need to ensure that the T- won't spread through one of the PromWds to another's right edge, as ALIGN(T-, L; PROMWD, R) might seem to allow. However, NOCROWD(T*BU) militates against this.

2.4. Example Derivation

Assuming that the tones involved are H*, L- and L%, (18) represents the prosodic structure of (5). Recall the constraints we have proposed thus far:

(17)
a. PROMWD=PROSWD
b. ALIGN-R(FOC; PROMWD)
c. T*BU=PROMWDSTRESSSYLL
d. ASSOC(T*BU, T*)
e. NOCROWD(T*BU)
f. ASSOC(T*, T*BU)
g. NOSPREAD(T*)
h. ASSOC(T-,]ip)
i. ALIGN-R(T-; ip)
j. ALIGN(T-, L; PROMWD, R)

Though it is difficult to represent graphically, the left and right edges of the L- and the L% are all at once aligned the right edge of: PromWd, Foc, ip and IP. In this way, they have no duration, and are only tone targets for interpolation.[7] Most importantly, none of our constraints are violated in (18).[8]

(18)

Words:	They	watched	PSY	CHO.	
IP:	(			$_{L\%}$)	
ip:	(			$_{L-}$)	
foc:	(			)	
PromWd:			(	$_{H*}$)	
*T*BU:*			( )		
Tones:			(H*)	(L-)(L%)	

[7] They are also available as triggers for other intonational processes, such as upstep and downstep.

[8] A note about this notation. "$_{T}$)" means that, underlyingly, a T marks this constituent. It does not say anything about the realization of these tone targets.

2.5. Yes/No Questions

As another example of these constraints at work, let us look at some Yes/No Questions (YNQs), which will also be the source of crucial data later in this paper. YNQs have the tonal properties laid out in (19).

(19) *Tonal Properties of YNQs*

a. L* marks the PromWd of a focus domain
b. H- marks material between the L* word and the right edge of the ip
c. H% marks the last syllable of the Intonation Phrase (IP)

An example is given in (20), where "legumes" is both the focus domain and the prominent word

(20)

Words:	Are	LE GUMES	a	good	source	of	vitamins?	
IP:	(							$_{H\%}$)
ip:	(							$_{H-}$)
foc:		(	)					
PromWd:		(	$_{L*}$)					
*T*BU:*		( )						
Tones:		(L*)		(	H-		)(H%)	

3. OCP Effects

3.1. The Phenomenon

Given our constraints thus far, if a PromWd is multiply-stressed, we would expect each of the stressed syllables to be T*BUs, and thus we should expect multiple tone targets within that PromWd. For that reason, given the YNQ in (21), with the focus domain being on "Alexander" ([ə̀ləgzǽndɚ]), we expect two L* targets – one on the secondarily stressed [ə̀], and another on the primarily stressed [zǽn].

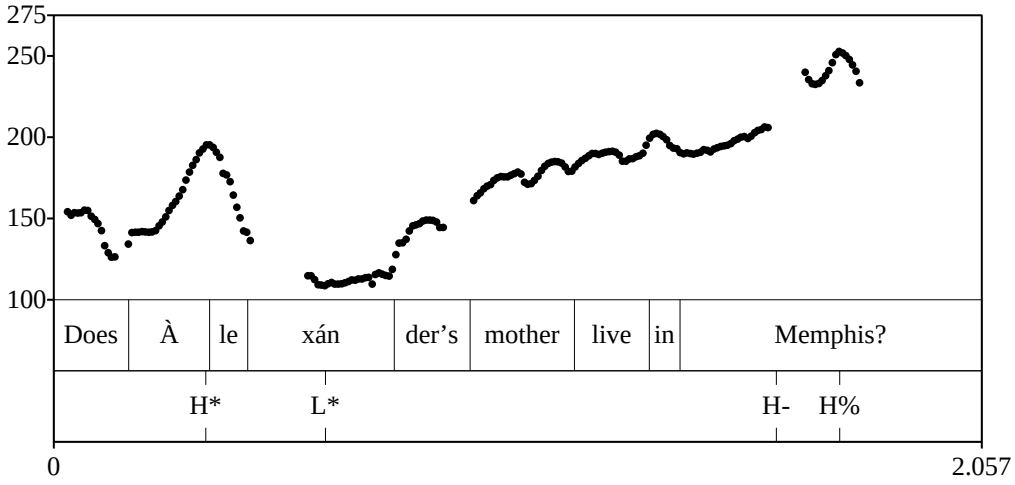
(21)

Words:	Does	A	LEXANDER'S	mother	live	in	Memphis?		
IP:	(							$_{H\%}$)	
ip:	(							$_{H-}$)	
foc:		(	)						
PromWd:		(	$_{L*}$)						
*T*BU:*		( )	( )						
Tones:		(L*)	(L*)	(	H-		)(H%)		

However, the representation in (21) is not necessarily what we find on the surface. Instead, we find that one T* changes into another. In the case of (21), the L* on [ə̀] surfaces as a H*. A pitch track of this surface representation is presented in (22).[9]

[9]Note that the H* is realized at the end of the [ə̀] syllable. This is because the timing of H* is such that it always realized at the end of the syllable it is linked to. It is therefore impossible to interpret this H* as linked to the [lə] syllable, as you might expect if this were a bitonal H+L*.

(22) *Example of the OCP Effect*



Given an example like the one in (23), which might be uttered after saying "And to my surprise…", this phenomenon seems to be limited in its domain.[10]

(23)

Words:	There	was	no	more	rice!	
IP:	(				$_{L\%}$)	
ip:	(				$_{L-}$)	
foc:	(				)	
PromWd:		( )(	)(	)(	)	
*T*BU:*		( )(	)(	)(	)	
Tones:		(H*)(	L*)(	L*	)(H*)(L-)(L%)	

Note that there are adjacent L*s, but there is no OCP effect. From this, it seems that we must posit a smaller domain of the OCP effect. It seems that the only boundary separating the two L*s in (23) that did not exist in (21) is that of a PromWd. For this reason, I believe that the domain of the OCP effect is the PromWd.[11]

3.2. Some More Data

As a pilot study, data was gathered from a small number of native speakers, who were given contexts such that multiply-stressed words were the focus of the YNQ sentence that they were recording. The data was collected using four- and five-syllable words that had various patterns of stress. These stress patterns and their respective words are given in (24).

[10] As should be clear from the representation, there can be multiple PromWds in a single focus domain, as long as one of their right edges aligns with the right edge of the focus domain. Cf. footnote 5.

[11] This is going to be equivalent to saying that the domain is the prosodic word, owing to PROMWD=PROSWD.

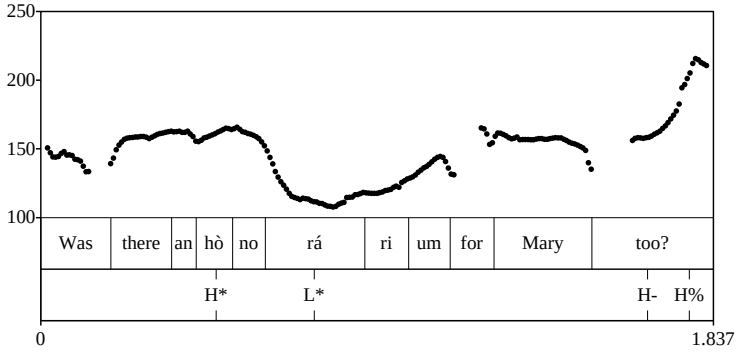
(24) *Tested Stress Patterns*[12]

a.	ˈσσˌσσ:	mó-de-rà-tor, á-lli-gà-tor
b.	σˈσσˌσ:	in-té-rro-gàte, o-ráng-u-tàn
c.	ˈσσˌσσσ:	ál-co-hò-lis-m, á-ma-tèu-ris-m
d.	σˈσσˌσσ:	ob-sér-va-tò-ry, vo-cá-bu-là-ry
e.	ˌσσˈσσ:	à-le-xán-der, mà-ri-juá-na
f.	σˌσσˈσ:	e-lèc-tio-neér, e-và-cu-ée
g.	ˌσσˈσσσ:	hò-no-rá-ri-um, nèu-ro-ló-gi-cal
h.	σˌσσˈσσ:	tu-bèr-cu-ló-sis, a-llì-te-rá-tion

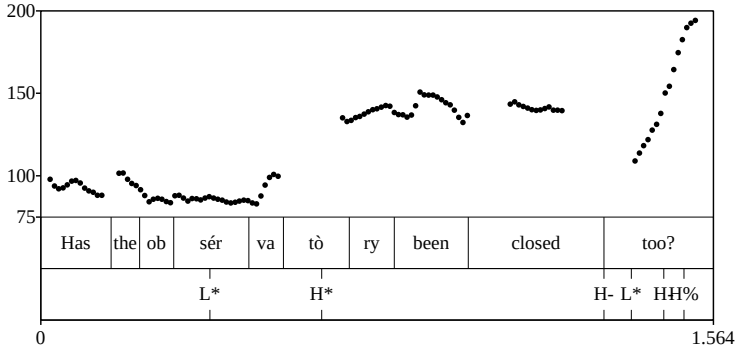
Below in (25) and (26), I give example pitch tracks for some of these. Pitch tracks for all of the words can be found in the Appendix.

It is important to note that there is no strict ordering of the [H*] and [L*]. Instead, what is relevant for their ordering is stress. The primary stress syllable carries the L*, and the secondary stress syllable carries the H*.

(25) *Hónoràrium*



(26) *Obsérvatòry*



[12]I should say that (24b) yielded no result, as it is an example of phonetic neutralization. The H* target on the final syllable would be timed simultaneously with the H target of the H-'s left edge.

3.3. The OCP Constraint Set

To account for the OCP effects that I have demonstrated, we need to incorporate a few more constraints into our grammar.

(27) OCP(T*, DOMAIN: PROMWD)
*Disallow identical T*s within the same PromWd. Assign a violation for every instance of multiple identical T*s within the same PromWd.*

(28) IDENT(T*-ˈσ)
Assign a violation if you change the tonal identity of T linked to a primary stress syllable.*

(29) IDENT(T*-ˌσ)
Assign a violation if you change the tonal identity of T linked to a secondary stress syllable.*

(30) MAX(T)
Do not delete tones from that the grammar calls for. Assign a violation for each deleted T.

(31) MAX(ASSOC)
Do not delete association lines. Assign a violation every time an association line is deleted – even if it is reassociated elsewhere.

(32) DEP(T)
Do not epenthesize any Ts. Assign a violation for the epenthesis of a T.

The crucial ranking here is that IDENT(T*-ˌσ) is ranked below all of the other constraints. This is because the identity of a T* linked to a secondary stress syllable must be more violable than the identity of a T* linked to the primary stress syllable in order to generate the correct output. In fact, this is the only constraint we ever violate in our entire system, which is the same reason that we are able to rank it.

Furthermore, this ranking addresses the questions, "why change the T*'s identity? Why not simply delete it?" The fact that the MAX constraints outrank the IDENT(T*-ˌσ) constraint is enough. Moreover, the DEP constraint prevents the insertion of frivolous tones that would break up the OCP effect. This of course means that the DEP constraint must also outrank IDENT(T*-ˌσ).

We have addressed the phenomenon of the change, but we have yet to address the phenomenon of L* to H*. It seems to me that, given an inventory where there are the five pitch accents given in (1), the only possible change (without changing to a bitone, which would involve violating DEP) is from L* to H*, and vice-versa.[13]

3.4. Other Remarks

Some possible supplemental evidence for this H* as a phonetic realization of L* is that naïve speakers often do not perceive this H* as existing, even if they are told to listen for it – instead,

[13] One example that I elicited seems to have a mid-range target on the secondary stress syllable, where there should be an underlying L*. I believe it to be M* as there is flat interpolation from the beginning of the sentence to the M*, and that plateau is significantly lower than the speaker's high range, and significantly higher than his low range.

they hear L*. This is can easily be accounted for, assuming that this [H* L*] sequence is essentially a phonetic variant of [L* L*].

Just as many naïve speakers of English cannot perceive the distinction between [p] and [ph], as there are no meaningful contrasts, they cannot perceive [H* L*] as distinct from [L* L*]. Anecdotally, it should be mentioned that speakers who do this, are often unaware of it; and, when asked to produce [L* L*], cannot necessarily do so on the spot. [14]

3.5. H* H* OCP Effects

There are similar effects with H*s. However, previous research that may have noticed these could have gone unnoticed, as they would able to be modeled in standard ToBI as just a L+H* bitone. Shilman (2006) found that one of the few differences between American English Motherese and standard American English was "the increased use of L+H* pitch accents relative to H* pitch accents" in Motherese. Perhaps what she found was not an increased change in underlying pitch accent, but rather in phonetic realization of pitch accent.

That said, there should be some differences between L+H* and H* with an OCP L. We should expect that the OCP L would prefer secondary stress in the same way that the OCP H* does. Also, we should expect there to be no strict ordering between the L and the H; instead, we should find L following secondary stress. Data such as that in Shilman must be returned to with the possibility of L+H* not being L+H* underlyingly.

4. Future Research

There is much research that is left open as a result of this short study.

First and foremost, the general constraint set needs to be refined in many ways. Specifically, there need to be accounts for at least: boundary tones, boundary tone/phrase accent interaction (upstep, downstep), pitch accent timing (H* at the end of a syllable, L* throughout the syllable, bitones).

Second, it seems that this is OCP process is optional – even in dialects that do it rather often. What are the other optimal candidates? Do they involve epenthesis, deletion, or no change at all? And in what way can we modify the grammar proposed in this paper to produce all the surfacing candidates?

Third, and most instersting of all, it seems that there are times when the H is not necessarily linked to a secondary stress syllable. In fact, there are two examples that quickly come to mind where there is a HL sequence other than the kind described here.[15] In one, the H is on the immediately proceeding syllable, even if not stressed. This looks to be a good candidate for H+L*. An example is given below.

(33) Does MARIA's mother live in Memphis?
H L* (H-)(H%)

In the other, there is a high plateau before the L*, covering all syllables from non-prominent

[14] At its weakest, this supplemental "evidence" favors any analysis whereby this HL sequence is a phonetic variant of a LL sequence. However, in any other analysis, it would be difficult to explain (a) the stressed-syllable timing of the H, or (b) the answer to "why a target whose value is opposite of what you expect?"

[15] Special thanks to Kie Zuraw, for pointing this out to me.

words. This might be a candidate for -H – a phrase accent that associates to the left edge of the ip. An example is given below.

(34) Do you have to do that NOW?
(-H) L* (H-)(H%)

What is so interesting about this family of HL* sequences is that they seem to possibly share some properties. Namely that they are allophonic with respect to L*, and they seem to all give the feeling that the speaker is trying to be more engaging. These properties would also be consistent with Shilman's result that Motherese contains L+H* where one might otherwise find a simple H*. That is, this OCP effect might be the result of a register shift or it may mark enthusiasm, an attempt at extra salience, or something similar.

As a follow-up to the HL* family, we should also expect a possible L*H family as well, if we reverse primary and secondary stress. Moreover, we might also expect H*L and LH* families as a result of OCP effects on [H* H*] sequences.

5. Conclusion

This paper makes the case for an OT grammar of intonation, the constraints of which are rarely violated. This paper also shows that there are OCP effects in this grammar, with special regard for [L* L*] sequences becoming [H* L*] or [L* H*], depending on stress pattern.

Furthermore, it would be difficult to explain this data using bitones for multiple reasons. First, bitones do not exhibit any attraction to secondary stress. Second, and more importantly, we have [L* H*] as well as [H* L*], in the same environment – this would mean, if we were to use bitones, we would need L*+H as well as H+L*, where the H is attracted to secondary stress syllables. This seems uncharactaristically complicated for the intonational grammar, whose tones otherwise have a simple attraction to stress. My OCP account for these patterns is both strong, making clear predictions, and flexible, as its exact conditioning environment is open to future work.

References

Beckman, Mary E., and Gayle M. Ayers. 1997. Guidelines for ToBI labelling. *The OSU Research Foundation* 3. Available at http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hayes, Bruce, and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9:47–96.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, ed. Philip R. Cohen, Jerry L. Morgan, and Martha E. Pollack, 271–311. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1996. The prosodic structure of function words. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James Morgan and Katherine Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.

Shilman, Molly. 2006. Intonational phonology of American English Motherese. UCLA., Mar 2006.

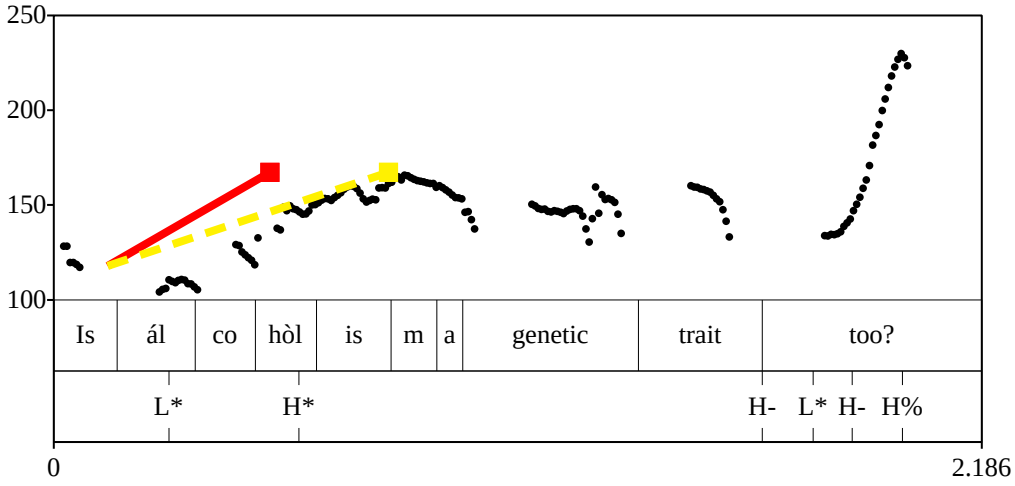
Silverman, Kim, Mary Beckman, John Pitrelli, Mari Ostendorf, Colin Wightman, Patti Prince, Janet Pierrehumbert, and Julia Hirschberg. 1992. TOBI: A standard for labeling English prosody. In *Second International Conference on Spoken Language Processing*.

Welby, Pauline. 2003. Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46:53–81.

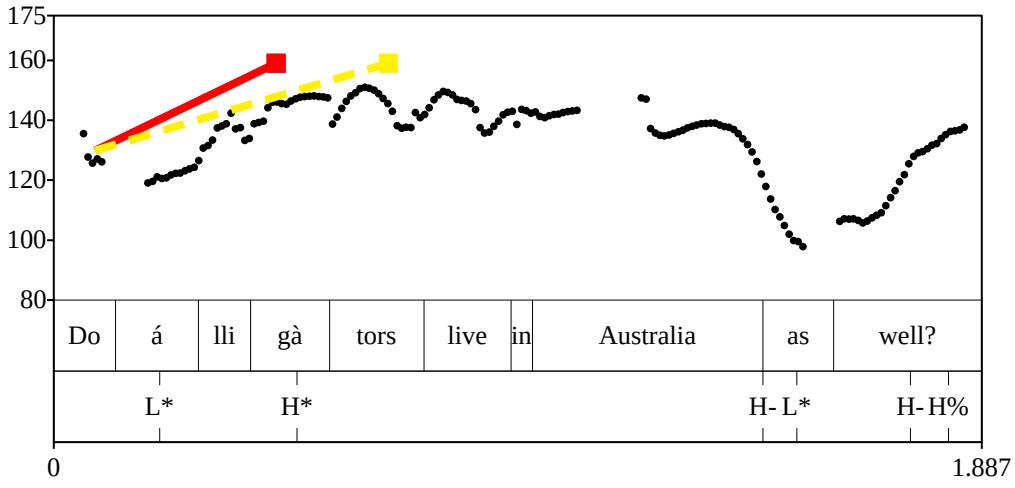
Appendix

Throughout this Appendix, the solid line indicates the actual path of interpolation. Conversely, the dotted line path indicates the path of interpolation had there been no H* arising from the OCP Effect.

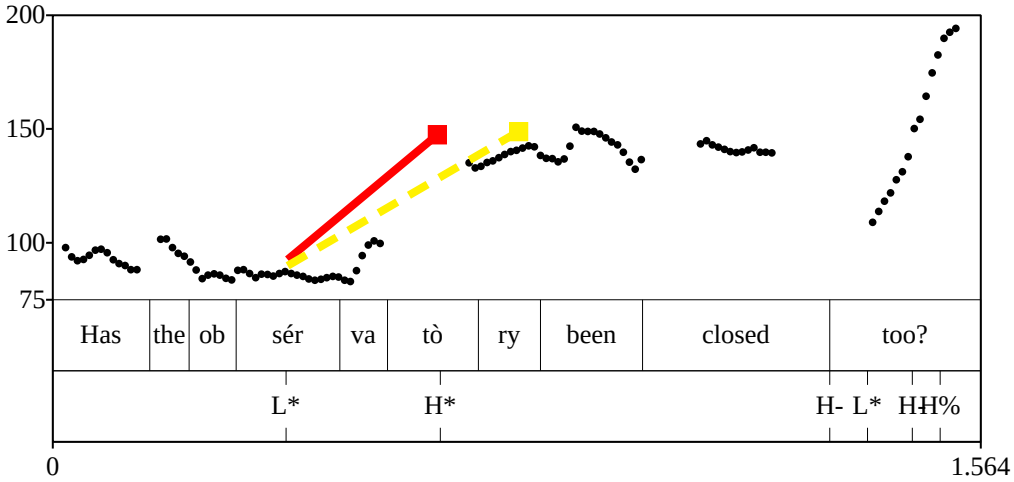
(35) *Álcohòlism*



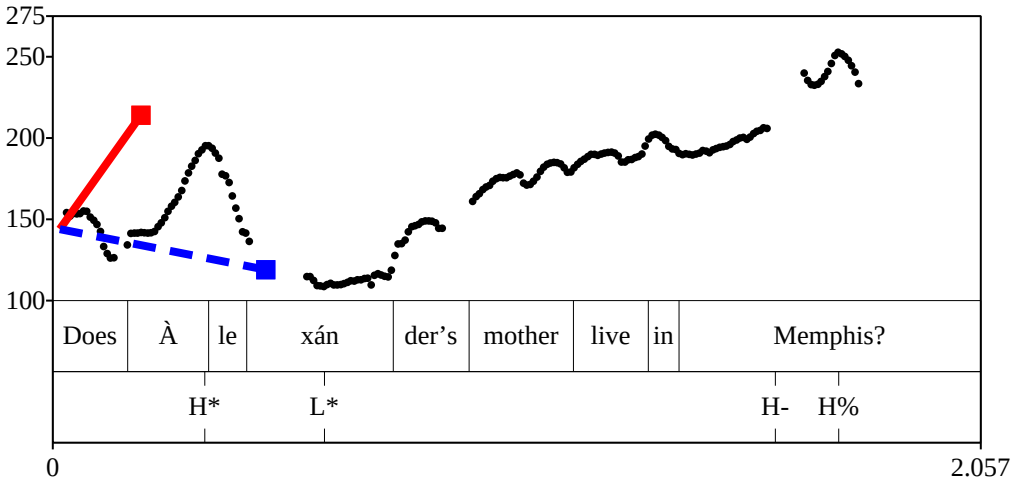
(36) *Álligàtors*



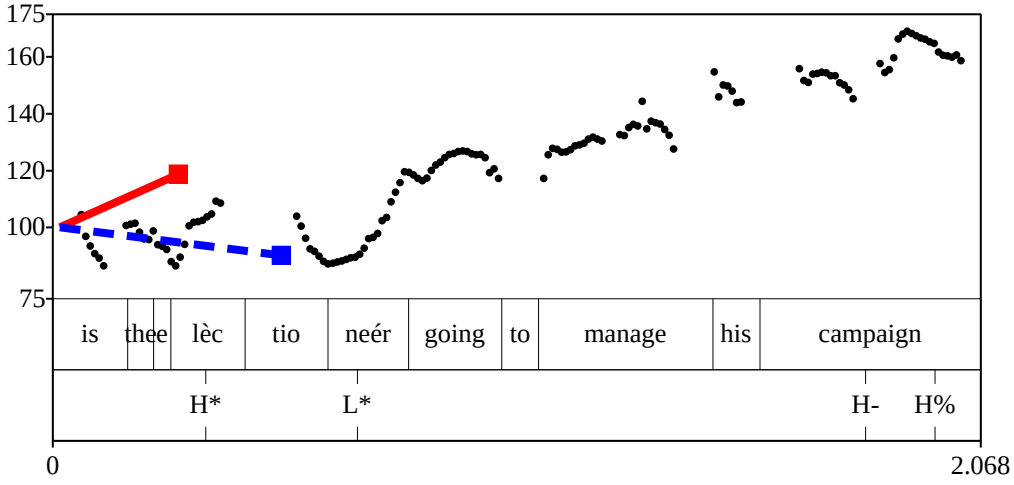
(37) *Obsérvatòry*



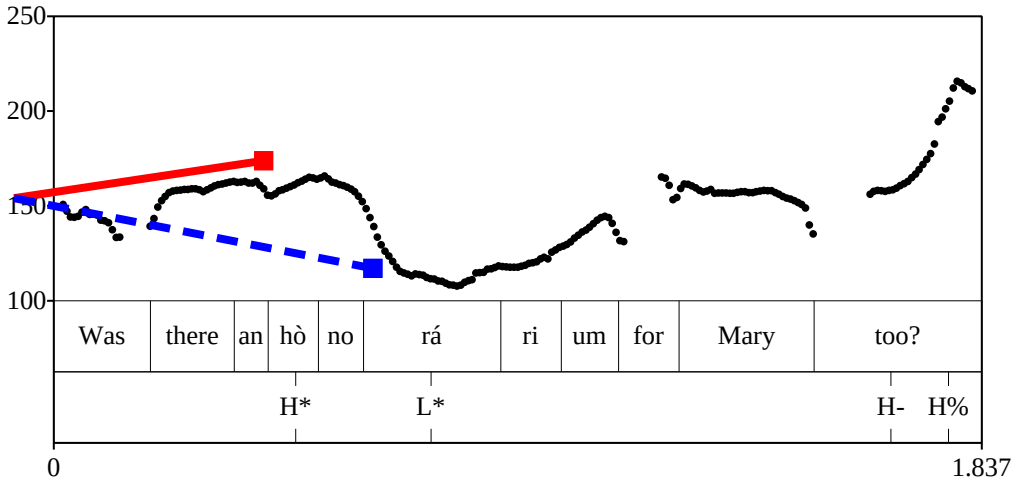
(38) *Àlexánder*



(39) *Elèctionèer*



(40) *Hónoràrium*



(41) *Tubèrculósis*