(23) FIRST RELAXATION:

minimize: $\langle \mathbf{\Sigma}_{n,m}, \mathbf{X} \rangle$;

subject to: $\mathbf{X} \in \mathcal{P}^n_{\text{rel}}$ s.t. $\langle \mathbf{A}_{\mathbf{e}}, \mathbf{X} \rangle \leq 0$ for any ERC $\mathbf{e}$ of the ERC matrix.

(24) SECOND RELAXATION:

minimize: $\langle \mathbf{\Sigma}_{n,m}, \mathbf{X} \rangle$;

subject to: $\mathbf{X} \in \mathcal{P}^n_{\text{rel}}$ s.t. $\langle \mathbf{A}^{\bar{\jmath}}_{\mathbf{e}}, \mathbf{X} \rangle \geq 0$ for any ERC $\mathbf{e}$ of the ERC matrix and any stratum $\bar{\jmath} \in \{1, \ldots, n\}$.

These linear programs (23) and (24) are the *relaxations* of the two integer programs (18) and (20).

The relaxation of an integer program provides a lower bound on the solution of that integer program. This lower bound is used by solution algorithms for the integer program. Of course, linear relaxations that provide tight bounds yield improved solution algorithms for the original integer problem (Bertsimas and Weismantel, 2005). Despite the fact that the two original integer programs (18) and (20) are equivalent, the two corresponding relaxations (23) and (24) are not. Claim 4 ensures that the feasible set of the relaxation (24) is a subset of that of the relaxation (23), so that the lower bound provided by a solution of the former will be at least as tight as the lower bound provided by a solution of the latter.

Claim 4 *If a matrix* $\mathbf{X}$ *belongs to the feasible set of problem (24), then it also belongs to the feasible set of problem (23).* ■

The following counterexample shows that the lower bound provided by the relaxation (24) is not just as tight as but actually tighter than the bound provided by the relaxation (23). Given the ERC matrix (25), the solution to the corresponding problem (7) is the ranking $F_2 \gg M \gg F_1$: the faithfulness constraint F_1 is redundant and should therefore be ranked at the bottom.

(25)
$$\mathbf{E} = \begin{array}{c} \begin{array}{ccc} F_1 & F_2 & M \end{array} \\ \left[\begin{array}{ccc} \text{W} & \text{W} & \text{L} \\ e & \text{W} & \text{L} \end{array} \right] \end{array}$$

The solutions of the two corresponding relaxations (23) and (24) are provided in (26).[3]

(26)
$$\mathbf{X}_{(23)} = \begin{array}{c} \\ F_1 \\ F_2 \\ M \end{array} \begin{array}{c} \begin{array}{ccc} st1 & st2 & st3 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{array} \right] \end{array} \quad \mathbf{X}_{(24)} = \begin{array}{c} \\ F_1 \\ F_2 \\ M \end{array} \begin{array}{c} \begin{array}{ccc} st1 & st2 & st3 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right] \end{array}$$

The relaxation (23) has a non-integral solution; the relaxation (24) is thus stronger because its solution is integral. The latter solution indeed represents the desired ranking, as it assigns F_2 to the top 3rd stratum (because of the 1 in the second column and third row) and F_1 to the bottom 1st stratum (because of the 1 in the first row and first column).

4 Conclusion

In this paper, I have focused on Prince and Tesar's (2004) formulation (7) of the problem of the acquisition of phonotactics, in terms of the alleged restrictiveness measure (9). This problem is NP-complete. To cope with this hardness result, in this paper I have looked for an integer programming formulation of the latter problem. The formulation in (20) has emerged as the best formulation among those considered, namely the one that yields the tightest relaxation. This problem (20) is an instance of a classical integer program, namely the *Assignment problem* with *linear side constraints* (AssignLSCsPbm). The result obtained in this paper thus paves the way for the efficient application of approximation algorithms for the AssignLSCsPbm to the problem of the acquisition of phonotactics in OT. In Magri (2012a), I report simulation results with Arora's et. al. (2002) algorithm, a state-of-the-art approximation algorithm for the AssignLSCsPbm.

Acknowledgments

I wish to thank A. Albright for endless discussion on the problem of the acquisition of phonotactics. This work was supported in part by a 'Euryi' grant from the European Science Foundation to P. Schlenker, by a grant from the Fyssen Research Foundation, and by the LABEX-EFL grant.

Appendix: proof of the main results

A.1 Proof of claim 1

Consider the example of the permutation matrix $\mathbf{X}$ in (27). There are seven constraints (hence $n = 7$), four of which are faithfulness constraints (hence $m = 4$). I have fringed each row of $\mathbf{X}$ with the name of the constraint it corresponds to and I have

[3] These solutions have been computed with the Matlab codes `RelaxedSubPbmFirstFormulation.m` and `RelaxedSubPbmSecondFormulation.m`, that solve the two relaxations (23) and (24), respectively. These codes are available on the author's website. The two codes use the two subroutines `MatrixToVectorConverter.m` and `VectorToMatrixConverter.m`, that are available on the author's website too.

fringed each column of $\mathbf{X}$ with the stratum it corresponds to.

$$\text{(27)}\quad \mathbf{X} = \begin{array}{cc} & \begin{array}{ccccccc} st1 & st2 & st3 & st4 & st5 & st6 & st7 \end{array} \\ \begin{array}{c} F_1 \\ F_2 \\ F_3 \\ F_4 \\ M_5 \\ M_6 \\ M_7 \end{array} & \left[\begin{array}{ccccccc} 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{array}\right] \end{array}$$

As prescribed by our conventions, the first four rows correspond to the four faithfulness constraints, the bottom three rows correspond to the markedness constraints; the leftmost column corresponds to the bottom stratum and the rightmost column corresponds to the top stratum.

The ranking $\pi_{\mathbf{X}}$ that corresponds to $\mathbf{X}$ can be obtained as follows: the 1 in the first column of $\mathbf{X}$ says that the markedness constraint M_6 is assigned by $\pi_{\mathbf{X}}$ to the bottom stratum $j = 1$; the 1 in the second column of $\mathbf{X}$ says that the faithfulness constraint F_1 is assigned to the next stratum $j = 2$; and so on. Thus, $\pi_{\mathbf{X}}$ is the ranking (28).

$$\text{(28)}\quad F_4 \gg F_2 \gg M_7 \gg M_5 \gg F_3 \gg F_1 \gg M_6$$

According to (9), the restrictiveness $\mu(\pi_{\mathbf{X}})$ of this ranking $\pi_{\mathbf{X}}$ is $8 = 3{+}3{+}1{+}1$: 3 markedness constraints underneath F_4, another 3 underneath F_2, 1 underneath and F_3 as all as underneath F_1. Here is a way to quickly compute this number directly from the permutation matrix $\mathbf{X}$.

Consider the matrix (29) obtained from the matrix (27) through the following two steps. *First*, all 1's which appear in the bottom three rows of $\mathbf{X}$ (and thus correspond to markedness constraints) are replaced with 0's.

$$\text{(29)}\quad \begin{array}{cc} & \begin{array}{ccccccc} st1 & st2 & st3 & st4 & st5 & st6 & st7 \end{array} \\ \begin{array}{c} F_1 \\ F_2 \\ F_3 \\ F_4 \\ M_1 \\ M_2 \\ M_3 \end{array} & \left[\begin{array}{ccccccc} 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{5} & 0 \\ 0 & 0 & \mathbf{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \end{array}$$

Second, each 1 which appears in one of the top four rows of $\mathbf{X}$ (and thus corresponds to a faithfulness constraint) is replaced with the number which identifes the corresponding column, diminished by 1. Thus for example, the 1 in the second row in the matrix $\mathbf{X}$ in (27) is replaced by a 5 in (29), since it occurs in the sixth column.

Next, let's scan the columns of the matrix (29) from left to right, assigning to each column which is not all zeros a progressive index k starting from $k = 0$, as made explicit in (30).

$$\text{(30)}\quad \begin{array}{cc} & \begin{array}{ccccccc} & k_1{=}0 & k_3{=}1 & & & k_2{=}2 & k_4{=}3 \end{array} \\ \begin{array}{c} F_1 \\ F_2 \\ F_3 \\ F_4 \\ M_1 \\ M_2 \\ M_3 \end{array} & \left[\begin{array}{ccccccc} 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{5} & 0 \\ 0 & 0 & \mathbf{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \end{array}$$

Now we can straightforwardly read out of (30) the number of markedness constraints ranked by $\pi_{\mathbf{X}}$ below each faithfulness constraint: F_1 has only one markedness constraint ranked below it, which is precisely the number $i_1 = 1$ which appears in the row corresponding to F_1 diminished by the value $k_1 = 0$ which corresponds to the column where that number appears; F_2 has three markedness constraints ranked below it, which is precisely the number $i_2 = 5$ which appears in the row corresponding to F_2 diminished by the value $k_2 = 2$ which corresponds to the column where that number appears; and so on.

Since $\mu(\pi_{\mathbf{X}})$ is defined in (9) as the sum over each faithfulness constraint of the number of markedness constraints ranked below that faithfulness constraint, we get the right result as in (31).

$$\text{(31)}\quad \begin{aligned} &\mu(\pi_{\mathbf{X}}) = \\ &\quad = \mu(F_1) + \mu(F_2) + \mu(F_3) + \mu(F_4) \\ &\quad = (i_1 - k_1) + (i_2 - k_2) + (i_3 - k_3) + (i_4 - k_4) \\ &\quad = (1-0) + (5-2) + (2-1) + (6-3) \\ &\quad = 8 \end{aligned}$$

Note that the sum in the second line of (31) can be rearranged as follows:

$$\text{(32)}\quad \begin{aligned} &\mu(\pi_{\mathbf{X}}) = \\ &\quad = (i_1 + i_2 + i_3 + i_4) - (k_1 + k_2 + k_3 + k_4) \\ &\quad = (i_1 + i_2 + i_3 + i_4) - (0 + 1 + 2 + 3) \end{aligned}$$

It is trivial to check directly from the definition (12) of scalar product that the first term $i_1 + i_2 + i_3 + i_4$ in the second line of (32) is the scalar product $\langle \mathbf{\Sigma}_{7,4}, \mathbf{X} \rangle$ between the permutation matrix $\mathbf{X}$ in (27) and the matrix $\mathbf{\Sigma}_{7,4}$ in (14). Thus, the first term in the second line of (32) corresponds to the first term in (15). It is also trivial to check that the second term $0+1+2+3$ in the second line of (32) is equal to $\frac{1}{2}m(m-1)$ for $m=4$. Thus, the second term in the second line of (32) corresponds to the second term in (15).

A.2 Proof of claim 2

Consider a ranking π, namely a permutation over $\{1,\ldots,n\}$. Let π^{-1} be its inverse. Recall that $\pi(i)=j$ means that constraint C_i is assigned by the ranking π to the jth stratum, with the top stratum being the one corresponding to $j=n$. Thus, $\pi^{-1}(j)$ is the constraint assigned by π to the jth stratum. Given an ERC $\mathbf{e}=[e_1,\ldots,e_n]$, let $k=k(\mathbf{e})\in\{1,\ldots,n\}$ be univocally defined by conditions (33): they say that the constraints assigned by π to the top strata $k+1,\ldots,n$ all have an e in the ERC $\mathbf{e}$ so that the constraint assigned by π to the kth stratum is the highest one that does not have an e in the ERC.

(33) a. $e_{\pi^{-1}(k+1)}=\ldots=e_{\pi^{-1}(n)}=e.$
b. $e_{\pi^{-1}(k)}\neq e.$

Thus, π is OT-consistent with the ERC $\mathbf{e}$ iff $e_{\pi^{-1}(k)}=\text{W}$. To prove Claim 2, I thus prove the equivalence (34), where $\mathbf{X}_\pi=[x_{i,j}]_{i,j=1}^n$ is the permutation matrix corresponding to π and $\mathbf{A}_\mathbf{e}=[a_{i,j}]_{i.j=1}^n$ is the matrix defined in (17).

(34) $\quad \langle\mathbf{A}_\mathbf{e},\mathbf{X}_\pi\rangle>0 \iff e_{\pi^{-1}(k)}=\text{W}.$

Assume that $e_{\pi^{-1}(k)}=\text{W}$; then I can reason as follows, following Prince and Smolensky (2004):

(35)
$$\begin{aligned}\langle\mathbf{A}_\mathbf{e},\mathbf{X}_\pi\rangle &= \sum_{i,j=1}^n x_{i,j}a_{i,j} &&= \sum_{i,j=1}^n x_{i,j}2^j t_i\\ &= \sum_{i=1}^n t_i\sum_{j=1}^n x_{i,j}2^j &&= \sum_{i=1}^n t_i 2^{\pi(i)}\\ &= \sum_{j=1}^n t_{\pi^{-1}(j)}2^j &&> 2^k-\sum_{j=1}^{k-1}2^j>0\end{aligned}$$

The proof of the reverse implication is analogous.

A.3 Proof of claim 3

To illustrate why claim 3 holds, consider the concrete case of the ERC $\mathbf{e}$ in (36).

(36)
$$\mathbf{t}=\begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & C_4 & C_5 \end{array}\\ \left[\begin{array}{ccccc} \text{W} & \text{W} & e & e & \text{L} \end{array}\right]\end{array}$$

A ranking π is OT-consistent with this ERC $\mathbf{e}$ provided it ranks either C_1 or C_2 above C_5. This condition is equivalent to the set of implications (37). For example, the the third implication says that if, π assigns C_5 to either stratum 3, or 4 or 5 (the latter being the top stratum), then π must assign either C_1 or C_2 to either stratum 4 or 5.

(37)
$$\begin{aligned} C_5\in\{1,2,3,4,5\} &\Longrightarrow C_1\in\{2,3,4,5\} \vee C_2\in\{2,3,4,5\}\\ C_5\in\{2,3,4,5\} &\Longrightarrow C_1\in\{3,4,5\} \vee C_2\in\{3,4,5\}\\ C_5\in\{3,4,5\} &\Longrightarrow C_1\in\{4,5\} \vee C_2\in\{4,5\}\\ C_5\in\{4,5\} &\Longrightarrow C_1\in\{5\} \vee C_2\in\{5\}\\ C_5\in\{5\} &\Longrightarrow C_1\in\emptyset \vee C_2\in\emptyset \end{aligned}$$

Consider the permutation matrix $\mathbf{X}=[x_{i,j}]_{i,j=1}^{n=5}$. Recall that $x_{i,j}=1$ iff the corresponding ranking π satisfies the condition $\pi(i)=j$ namely it assigns constraint C_i to the jth stratum. Thus, the implications in (37) can be restated in terms of permutation matrices rather than rankings as in (38), in the sense that a ranking π satisfies (37) iff the corresponding permutation matrix $\mathbf{X}_\pi$ satisfies (38). The five inequalities (38) can be written in

(38)
$$\begin{array}{rrrrrcrrrrrrrr} x_{5,1} & +x_{5,2} & +x_{5,3} & +x_{5,4} & +x_{5,5} & \leq & x_{1,2} & +x_{1,3} & +x_{1,4} & +x_{1,5} & +x_{2,2} & +x_{2,3} & +x_{2,4} & +x_{2,5}\\ & x_{5,2} & +x_{5,3} & +x_{5,4} & +x_{5,5} & \leq & & x_{1,3} & +x_{1,4} & +x_{1,5} & & +x_{2,3} & +x_{2,4} & +x_{2,5}\\ & & x_{5,3} & +x_{5,4} & +x_{5,5} & \leq & & & x_{1,4} & +x_{1,5} & & & +x_{2,4} & +x_{2,5}\\ & & & x_{5,4} & +x_{5,5} & \leq & & & & x_{1,5} & & & & +x_{2,5}\\ & & & & x_{5,5} & \leq & 0 \end{array}$$

(39) $2x_{5,1}+4x_{5,2}+8x_{5,3}+16x_{5,4}\leq 2x_{1,1}+4x_{1,2}+8x_{1,3}+16x_{1,4}+32x_{1,5}+2x_{2,1}+4x_{2,2}+8x_{2,3}+16x_{2,4}+32x_{2,5}$

(40)
$$\begin{array}{rrrrcrrrrrrrr} 4x_{5,1} & +4x_{5,2} & +4x_{5,3} & +4x_{5,4} & \leq & 4x_{1,2} & +4x_{1,3} & +4x_{1,4} & +4x_{1,5} & +4x_{2,2} & +4x_{2,3} & +4x_{2,4} & +4x_{2,5}\\ & 4x_{5,2} & +4x_{5,3} & +4x_{5,4} & \leq & & 4x_{1,3} & +4x_{1,4} & +4x_{1,5} & & +4x_{2,3} & +4x_{2,4} & +4x_{2,5}\\ & & 8x_{5,3} & +8x_{5,4} & \leq & & & 8x_{1,4} & +8x_{1,5} & & & +8x_{2,4} & +8x_{2,5}\\ & & & 16x_{5,4} & \leq & & & & 16x_{1,5} & & & & +16x_{2,5} \end{array}$$

(41) $4x_{5,1}+8x_{5,2}+16x_{5,3}+32x_{5,4}\leq 4x_{1,2}+8x_{1,3}+16x_{1,4}+32x_{1,5}+4x_{2,2}+8x_{2,3}+16x_{2,4}+32x_{2,5}$