OCP and Line-Crossing Constraint as aids to word segmentation

Paul Boersma, University of Amsterdam

Recently, Boll-Avetisyan and Kager (2014, henceforth B&K) showed that Dutch listeners can use labiality as a word segmentation cue in pre-lexical perception. B&K claim, without providing any explicit formalization, that Boersma's (1998, 2000, 2009) version of the Obligatory Contour Principle (OCP) cannot handle the experimental facts. In this talk I show that B&K's pre-lexical perception results are in fact straightforwardly accounted for with Boersma's autosegmental constraint sets, without any additional assumptions.

In Optimality Theory, Myers (1997) formulated the OCP as a violable constraint on phonological surface representations. Boersma (1998, 2000, 2009) argued that the Line-Crossing Constraint (LCC), too, should be regarded as a violable constraint on phonological surface representations, and that both OCP and LCC should be parametrized by the amount of phonological material that intervenes between the two elements:

(1) Autosegmental constraints (copied from Boersma 2009)

a. OCP (*feature value*, *material*): the surface form cannot contain two instances of *feature value* if not more than a certain amount of *material* intervenes;

b. LCC (*feature value*, *material*): a single instance of *feature value* in the surface form cannot span across a certain amount of *material*.

Boersma (1998, 2000) argued that these constraints are universally ranked by the amount of intervening material: OCP is ranked lower if there is more intervening material (e.g. lower across a prosodic word boundary than across a syllable boundary), whereas LCC is ranked higher if there is more intervening material (e.g. higher across a prosodic word boundary than across a syllable boundary). To explain a widely attested ban on having two labials within the same word, the labial versions of both constraints have to be high-ranked (the input is an underlying form, the candidates are surface forms; the indexes denote labial specifications):

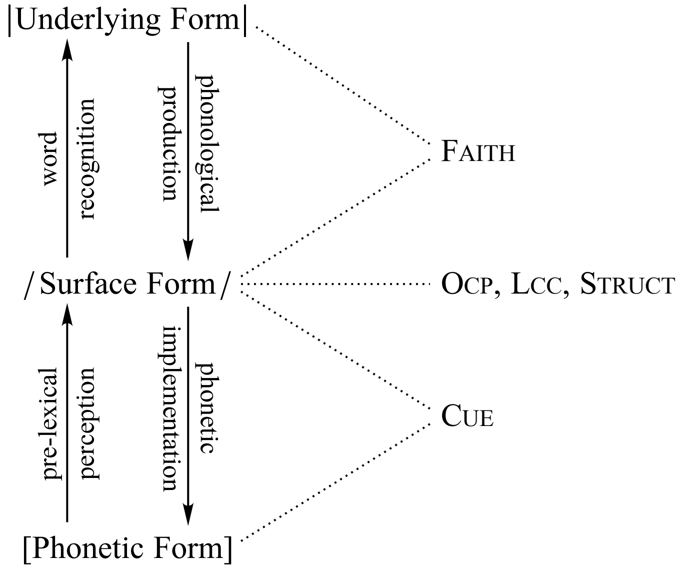
(2) Deleting one of the two underlying labials

$\|\{pama\}_{Wd}\|$	LCC (*lab*, $\}_{Wd}\{$)	OCP (*lab*, $\}_{\sigma}\{$)	LCC (*lab*, $\}_{\sigma}\{$)	IDENT (PLACE)	OCP (*lab*, $\}_{Wd}\{$)
$/\{\{ p_1 a \}_\sigma \{ m_1 a \}_\sigma \}_{Wd} /$			*!		
$/\{\{ p_1 a \}_\sigma \{ m_2 a \}_\sigma \}_{Wd} /$		*!			
☞ $/\{\{ p_1 a \}_\sigma \{ n a \}_\sigma \}_{Wd} /$				*	

We see here that if (and only if) both OCP and LCC are ranked above faithfulness, the surface form deletes one of the underlying labials. Given the simplified underlying form and constraint set employed here, this result is independent of whether the underlying |p| and |m| share their labial specification or not.

To see how Boersma's versions of OCP and LCC affect pre-lexical perception, we have to realize that they are part of a multi-level model of phonology and phonetics, in which they influence the output of phonological production, as in (2), as well as the output of pre-lexical perception, as later in tableau (4):

(3) How Ocp and Lcc influence production and comprehension



In the following pre-lexical perception tableau, the input is the auditory-phonetic form, and the output candidates are phonological surface forms that differ in the number of perceived words and in the number of labial features (Cue(place) is a cue constraint that punishes the perception of auditory labiality as a coronal feature value):

(4) Perceiving a word boundary

[pama]	Lcc (*lab*, $\}_{Wd}\{\,$)	Ocp (*lab*, $\}_{\sigma}\{\,$)	Lcc (*lab*, $\}_{\sigma}\{\,$)	Cue (place)	Ocp (*lab*, $\}_{Wd}\{\,$)
$/\{\{p_1\ a\}_\sigma\{m_1\ a\}_\sigma\}_{Wd}/$			*!		
$/\{\{p_1\ a\}_\sigma\{m_2\ a\}_\sigma\}_{Wd}/$		*!			
$/\{\{p_1\ a\}_\sigma\{n\ a\}_\sigma\}_{Wd}/$				*!	
☞ $/\{\{p_1\ a\}_\sigma\}_{Wd}\{\{m_2\ a\}_\sigma\}_{Wd}/$					*
$/\{\{p_1\ a\}_\sigma\}_{Wd}\{\{m_1\ a\}_\sigma\}_{Wd}/$	*!				

The result is that that the listener, given two auditory labial segments, perceives two words. Together, the labial Ocp and the labial Lcc have conspired to help the listener segment the speech stream into words, which is precisely what B&K found in their Dutch listeners.

The talk contains much more detail than this. For instance, learning simulations with Dutch data show that the word segmentation effect of (4) happens only for labials and not for coronals, which have a much larger probability of co-occurrence within a Dutch word. Again, this is precisely what B&K attested in their Dutch listeners.

Boersma, P. (1998). *Functional phonologs*. PhD dissertation, University of Amsterdam.
Boersma, P. (2000). *The OCP in the perception grammar*. Ms. [Rutgers Optimality Archive **435**]
Boersma, P. (2009). Cue constraints and their interactions in phonological perception and production. In P. Boersma & S. Hamann (eds.), *Phonology in perception*. Berlin: Mouton De Gruyter. 55–110.
Boll-Avetisyan, N., & R. Kager (2014). OCP-place in speech segmentation. *Language & Speech* **57**: 394–421.
Myers, J. S. (1997). OCP effects in Optimality Theory. *NLLT* **15**: 847–892.