Variability, cumulativity, and trigger asymmetries in Finnish.

Wendell Kimper
wkimper@ucsc.edu

University of California, Santa Cruz

January 30th, 2012 :: Phorum, UC Berkeley
In this talk, I argue that the choice between transparency and opacity in vowel harmony is best viewed as a competition between potential harmony triggers.

This argument is supported by evidence from a nonce-word experiment on Finnish disharmonic loanwords.

Preview of Coming Attractions:
- Transparency, opacity, and trigger competition
- Harmony and disharmony in Finnish loanwords
- The experiment
- Formal analysis: SHG
A Question and a Proposal

The generalization:
- Segments which fail to participate in vowel harmony can be either transparent (skipped by harmony) or opaque (blocking further propagation harmony).
- In languages with multiple non-participants, some segments may be transparent, while others are opaque.

Question: What determines whether a segment is likely to be transparent or opaque?

Proposal: The choice between transparency and opacity represents a competition between potential harmony triggers, and is influenced by the relative strength of each trigger.

See e.g. (Hayes et al., 2009) for an application of this overall approach to transparency and opacity in Hungarian vowel harmony.
Trigger Competition

When $V_2$ is unable to undergo harmony, $V_1$ and $V_2$ must compete to determine the feature specification of $V_3$.

**Transparency**

\[
[+][-]
\]

$V_1$ $V_2$ $V_3$

Transparency represents a victory for $V_1$, while opacity represents a victory for $V_2$.

The outcome is determined by the relative importance of opposing influences on trigger strength:

- **Feature Dominance**: $V_1$ bears the dominant feature value, and dominant triggers are better triggers.

- **Locality**: $V_2$ is local to $V_3$, and local triggers are better triggers.
**Prediction:** transparency should be more likely *from* a strong trigger, and more likely *across* a weak trigger. So what inherent properties of segments make them better triggers?

**In front/back harmony, low vowels are better triggers.**

- Vowel harmony confers a perceptual advantage, because it boosts the salience of weak contrasts (Suomi, 1983; Kaun, 1995; Gallagher, 2010; Kimper, to appear).

- Lower vowels are poorly cued for a front/back contrast, and therefore have more incentive to spread.

Lower vowels should be more likely to induce transparent front/back harmony, but less likely to be transparent themselves.
Finnish Vowel Harmony: Native Words

Finnish vowel harmony: root vowels must agree in backness, and suffixes alternate to take on the specification of their roots.

Unpaired [i] and [e] do not undergo harmony. They are transparent — harmony skips them to affect other segments.

**Finnish Vowel Inventory**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>y</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>ø</td>
<td>o</td>
<td></td>
</tr>
<tr>
<td>æ</td>
<td></td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

- a. põytæ-næ ‘table-ESS’
- b. pouta-na ‘fine weather-ESS’
- c. tunte-vat ‘feel-3PL’
- d. ukit-han ‘grandfathers-EMPH’

For more on Finnish vowel harmony, see e.g. (Kiparsky, 1973, 1981; Goldsmith, 1985; Ringen, 1988; Steriade, 1987; Vago, 1988, and others).
Harmony and Disharmony in Loanwords

Disharmony also results in loanwords, where other front vowels [æ,ø,y] also fail to undergo harmony (Campbell, 1980; Ringen and Heinämäki, 1999).

Loanwords still induce harmonic alternations on suffixes. All-back stems take back suffixes, All-front stems take front suffixes (not shown), and disharmonic front–back stems take back suffixes.

- a. vulgæ:ri ‘vulgar’
- b. tyranni ‘tyrant’
- c. analy:si ‘analysis’
- d. afæ:ri ‘affair’
- e. rotunda-sta ‘rotunda’ (*rotunda-stæ)
- f. saluuna-sta ‘saloon’ (*saluuna-stæ)
- g. symptomi-a ‘symptom’ (*symptomi-æ)
- h. syntaksi-a ‘syntax’ (*syntaksi-æ)
Harmony and Disharmony in Loanwords

In back–front stems, speakers exhibit variation: either a back suffix (transparency) or a front suffix (opacity) is possible. (Neutral vowels [i] and [e] remain consistently transparent, as in the native phonology.)

a. hieroglyfi-a ~ hieroglyfi-æ ‘hieroglyph’
b. analy:si-a ~ analy:si-æ ‘analysis’
c. martty:ri-a ~ martty:ri-æ ‘martyr’
d. sutenø:ri-a ~ sutenø:ri-æ ‘pimp’
e. jonglø:ri-a ~ jonglø:ri-æ ‘juggler’
f. amatø:ri-a ~ amatø:ri-æ ‘amateur’
g. miljonæ:ri-a ~ miljonæ:ri-æ ‘millionaire’
h. afæ:ri-a ~ afæ:ri-æ ‘affair’

Variability, cumulativity, and trigger asymmetries in Finnish.
Ringen and Heinämäki (1999) conducted a study of 50 native speakers of Finnish, eliciting suffixed forms for back–front disharmonic loanwords. Subjects’ likelihood of treating a vowel as variably transparent increased with its height. Hayes and Londe (2006); Hayes et al. (2009) found a similar asymmetry in Hungarian transparency/opacity.

Front vowels [æ,ø,y] are not equally likely to be transparent.

<table>
<thead>
<tr>
<th>Height</th>
<th>Mean (x̄)</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>.35</td>
<td>hieroglyfi (.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>analy:si (.48)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>martty:ri (.32)</td>
</tr>
<tr>
<td>Mid</td>
<td>.18</td>
<td>sutenø:ri (.10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>jonglø:ri (.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>amatø:ri (.26)</td>
</tr>
<tr>
<td>Low</td>
<td>.12</td>
<td>miljonæ:ri (.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hydrosfæ:ri (.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>afæ:ri (.25)</td>
</tr>
</tbody>
</table>
The Experiment

A nonce-word experiment was conducted using novel back–front disharmonic loanwords; subjects were asked to choose between back suffixes (transparency) and front suffixes (opacity).

**Subjects**: 179 native speakers of Finnish, recruited using an online social networking site, meeting the following criteria:

- Responded to each of the items.
- Responded incorrectly on no more than one catch trial.

**Hypothesis**: The likelihood of transparency should increase when $V_1$ is lower, and decrease when $V_2$ is lower.

This experiment was designed in collaboration with Riikka Ylitalo (University of Oulu).
Stimuli

CV₁C.CV₂ nonce words; in all conditions, V₁ was back and V₂ was front (and non-neutral). Height and length of both V₁ ([a,o,u]) and V₂ ([æ,ø,y]) were manipulated orthogonally (36 total).*

Consonants [p,t,k,s,f] were used ([f] is not present in the native inventory but is permitted in loanwords). A randomly selected half of the medial CC sequences were geminates, and the remainder were heterorganic clusters.

A number of back–back and front–front words were include as catch trials; because they are fully predictable, errors indicate that a subject is not performing the task.

* One item (Co:C.Cø:) suffered from a typo (presented as Cø:C.Cø:) — it was analyzed as an additional catch trial.
Task

Subjects were presented with a frame sentence with a blank, and asked to make a binary choice between a nonce item with a front suffix and that same item with a back suffix.

Prosessissa tarvittava kemiallinen liuos valmistetaan ______

  a. puktyssa
  b. puktyssä

Frame sentences chosen centered thematically around chemistry (to encourage subjects to treat them as plausible loans).

Each nonce word was paired consistently with its own frame sentence, and all subjects saw each item once.
Results

A generalized linear mixed effects model was fitted on back responses, with random effects for subjects. Vowel height is treated as an ordered factor (orthogonal polynomial contrast coding).

**Significant Predictors:**

- $V_1$ height, linear ($|z|=1.617, p<0.001$).
- $V_2$ height, linear ($|z|=3.617, p<0.001$).
  - $V_2$ height, quadratic ($|z|=3.746, p<0.001$).
- $V_2$ length ($|z|=4.989, p<0.001$).

$V_1$ length was included, but not significant ($z=1.018, p=0.31$).

**Summary:** Transparency is more likely with a lower $V_1$, and more likely with a higher $V_2$ (length effects set aside for now).
Prop. Transparency by V1 Height

<table>
<thead>
<tr>
<th>Initial Vowel Height</th>
<th>Prop. Back Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>0.15</td>
</tr>
<tr>
<td>mid</td>
<td>0.20</td>
</tr>
<tr>
<td>low</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Variability, cumulativity, and trigger asymmetries in Finnish.
Summary of Results

The competing triggers approach predicts that segments which are good triggers will be more likely to induce transparency, and will be less likely to be transparent themselves.

Perceptually-impoverished low vowels are preferred triggers for front/back harmony. Therefore, transparency should be more likely when $V_1$ is low, and less likely when $V_2$ is low.

This prediction precisely matches Finnish speakers’ performance in the experimental task; $V_1$ and $V_2$ height had significant effects in the predicted direction.
Alternatives

Nevins (2004); Kaun (1995): less sonorous vowels are more likely to be transparent.
- Matches $V_2$ effects.
- Predicts no effect of $V_1$.

Benus and Gafos (2007): more articulatorily robust vowels are more likely to be transparent.
- Matches $V_2$ effects.
- Predicts either no $V_1$ effects, or opposite direction.

The competing triggers model is uniquely able to capture the role that $V_1$ plays in transparency/opacity in Finnish.
The Formal Model: Overview

I propose a formal account of vowel harmony in which:

- Harmony is driven by a positively defined constraint, which assigns rewards for spreading.
- A preference for locality is expressed by diminishing the reward received for harmony among non-adjacent elements.
- Asymmetries in triggering are expressed by increasing the reward received for harmony from a preferred trigger.

This account is situated in **Serial Harmonic Grammar (SHG)**, a variant of Optimality Theory (OT) which combines:

- The stepwise evaluation of Harmonic Serialism (HS)
- The weighted constraints of Harmonic Grammar (HG)

For more on Serial Harmonic Grammar, see Pater (2009); Mullin (2010).
The Constraint

**Spread**\((F)\): For a feature \(F\), assign a reward of \(+1\) for each segment linked to \(F\) as a dependent.

- **Scaling factor: non-locality**
  For a trigger \(\alpha\) and a target \(\beta\), multiply the reward earned for the dependent segment \(\beta\) by a constant \(k\) (such that \(1 > k > 0\)) for each unit of distance \(d\) intervening between \(\alpha\) and \(\beta\).

- **Scaling factor: perceptual salience**
  For a trigger \(\alpha\), a target \(\beta\), and a feature \(F\), multiply the reward earned for the dependent segment \(\beta\) by a constant \(x\) (such that \(x > 1\)) for each degree \(i\) to which \(\alpha\) is perceptually impoverished with respect to \(\pm F\).
In HG (Smolensky and Legendre, 2006), constraints are assigned weights rather than strict ranks. Optimality is determined by comparing candidates’ harmony scores:

\[ \mathcal{H}(A) = \sum_{i=1}^{n} w(C_i) \times C_i(A) \]

Multiple violations/rewards of lower-weighted constraints can “gang up” to override a higher-weighted constraint.

Scalar constraints can exploit this cumulativity to produce stringency effects (Pater, to appear).
Dominance and Locality: Psuedo-Finnish

Back is the dominant feature

- \( w(\text{SPREAD}(B)) > w(\text{SPREAD}(F)) \)

Spreading back non-locally ties with spreading Front locally

- \( w(\text{SPREAD}(B)) \times k = w(\text{SPREAD}(F)) \)

<table>
<thead>
<tr>
<th></th>
<th>( \text{SPREAD}(B) )</th>
<th>( \text{SPREAD}(F) )</th>
<th>( \mathcal{H} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>( uy \bar{\ddot{a}} )</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>b.</td>
<td>( uy \bar{a} )</td>
<td>+0.25 ( (+1 \times .25) )</td>
<td>+1</td>
</tr>
</tbody>
</table>

Variability, cumulativity, and trigger asymmetries in Finnish.
Quality Sensitivity, V1

If V1 is preferred as a trigger, it can overcome the disadvantage of spreading non-locally.

\[ w(\text{Spread}(B)) \times k \times x > w(\text{Spread}(F)) \]

<table>
<thead>
<tr>
<th></th>
<th>(\text{Spread}(B))</th>
<th>(\text{Spread}(F))</th>
<th>(J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>a. a (\ddot{y}) -(\ddot{a})</td>
<td>+1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>b. a (\ddot{y}) -a</td>
<td>+0.375 (+1 * .25 * 1.5)</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

Variability, cumulativity, and trigger asymmetries in Finnish.
If V2 is a preferred trigger, it can overcome the disadvantage of spreading a non-dominant feature.

\[ w(S\text{pread}(B)) \times k < w(S\text{pread}(F)) \times x \]

<table>
<thead>
<tr>
<th></th>
<th>(S\text{pread}(B)) (= 4)</th>
<th>(S\text{pread}(F)) (= 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>(u \ddot{a} - \ddot{a})</td>
<td>(+1.5) ((+1 \times 1.5))</td>
</tr>
<tr>
<td>b.</td>
<td>(u \dddot{a} - a)</td>
<td>(+0.25) ((+1 \times .25))</td>
</tr>
</tbody>
</table>
Modeling the Experimental Results

The values for the constants $k$ and $x$ in the scale factors can be derived from the data itself. From the mixed effects model for the Finnish results, focusing only on the high vs. low distinction:

- $k$ (locality) $= 0.25$
- $x$ (perceptual salience) $= 1.5$

Probabilistic data can be modeled using Maximum Entropy HG (Goldwater and Johnson, 2003; Jäger and Rosenbach, 2006), which defines a probability distribution across candidates based on the exponent of their harmony scores.

Experimental data were fed to a Maximum Entropy batch learner to find a set of weights that produces probabilities for each candidate closely matching attested proportions.
# V1 Effects in Finnish

<table>
<thead>
<tr>
<th>Word</th>
<th>$\text{Spread}(B)$</th>
<th>$\text{Spread}(F)$</th>
<th>$\mathcal{H}$</th>
<th>Obs.</th>
<th>MaxEnt</th>
</tr>
</thead>
<tbody>
<tr>
<td>pukty</td>
<td>3.297</td>
<td>1.924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. pukty-ssaä</td>
<td>+1</td>
<td>1.924</td>
<td>0.804</td>
<td>0.750</td>
<td></td>
</tr>
<tr>
<td>b. pukty-ssa</td>
<td>+0.25 ( +1 * .25 )</td>
<td>0.824</td>
<td>0.196</td>
<td>0.250</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Word</th>
<th>$\text{Spread}(B)$</th>
<th>$\text{Spread}(F)$</th>
<th>$\mathcal{H}$</th>
<th>Obs.</th>
<th>MaxEnt</th>
</tr>
</thead>
<tbody>
<tr>
<td>sakfy</td>
<td>3.297</td>
<td>1.924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. sakfy-stä</td>
<td>+1</td>
<td>1.924</td>
<td>0.687</td>
<td>0.665</td>
<td></td>
</tr>
<tr>
<td>b. sakfy-sta</td>
<td>+0.375 ( +1 * 1.5 * .25 )</td>
<td>1.236</td>
<td>0.313</td>
<td>0.335</td>
<td></td>
</tr>
</tbody>
</table>

Variability, cumulativity, and trigger asymmetries in Finnish.
## V2 Effects in Finnish

<table>
<thead>
<tr>
<th>sakfy</th>
<th>Spread(B)</th>
<th>Spread(F)</th>
<th>$H$</th>
<th>Obs.</th>
<th>MaxEnt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.297</td>
<td>1.924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. sakfy-stä</td>
<td></td>
<td>+1</td>
<td>1.924</td>
<td>0.687</td>
<td>0.665</td>
</tr>
<tr>
<td>b. sakfy-sta</td>
<td>$+0.375$</td>
<td>(+$1 \times 1.5 \times .25$)</td>
<td>1.236</td>
<td>0.313</td>
<td>0.335</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>tafsä</th>
<th>Spread(B)</th>
<th>Spread(F)</th>
<th>$H$</th>
<th>Obs.</th>
<th>MaxEnt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.297</td>
<td>1.924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. tafsä-ssä</td>
<td></td>
<td>$+1.36$</td>
<td>2.61</td>
<td>0.776</td>
<td>0.799</td>
</tr>
<tr>
<td>b. tafsä-ssa</td>
<td>$+0.375$</td>
<td>(+$1 \times 1.5 \times .25$)</td>
<td>1.236</td>
<td>0.223</td>
<td>0.200</td>
</tr>
</tbody>
</table>

Variability, cumulativity, and trigger asymmetries in Finnish.
Summary

The constraint driving harmony is positive and scalar

- Rewards for spreading are increased if the trigger is perceptually impoverished.
- Rewards for spreading are decreased if trigger and target are non-local.

Cumulative interaction of weighted constraints means that differences in the degree of reward received can usurp higher-weighted constraints.

Independently-motivated restrictions on locality and segmental triggers interact to produce quality-sensitive transparency/opacity asymmetries.
Central claim: In this talk, I have argued that the choice between transparency and opacity in vowel harmony represents a competition between potential triggers.

- This approach predicts that transparency should be more likely from a preferred harmony trigger, and less likely across one.

Evidence: This prediction is supported by the results from a nonce-word experiment testing variable transparency in Finnish disharmonic loanwords.

- Unlike alternative approaches, a trigger competition approach is able to explain both $V_2$ and $V_1$ effects.

Model: Trigger competition can be modeled using positive scalar harmony constraint in SHG.
Acknowledgments: Many thanks to John McCarthy, Joe Pater, and John Kingston for insightful discussion of this work. Experimental materials were designed in collaboration with Riikka Ylitalo. This research was supported in part by grant BCS-0813829 from the National Science Foundation to the University of Massachusetts, Amherst.
References I


References II


Kiparsky, Paul. 1981. Vowel harmony. MIT.


