Phonetic Explanation in Phonology: Appealing to Perceptibility Scales

Jason Bishop
University of Leipzig

Perceptibility scales have been appealed to in order to account for a number of phonological phenomena in various languages. This study examines the phonetic evidence for one scale in particular, that proposed by Jun (1995, 2004). First, Jun's proposal involving the perceptibility of /p,t,k/ and its significance to regressive place assimilation in consonant clusters is reviewed. Second, previous phonetic research on unreleased final /p,t,k/ is considered. It is argued that previous investigations into the perceptibility of unreleased /p,t,k/ –most of which have not supported Jun's (dorsal > labial > coronal) hierarchy– have not provided an accurate picture of their relative salience. Finally, a new perceptual study is carried out with controls for vowel environment and response bias. Results offer only little if any support for a (dorsal > labial > coronal) hierarchy, but raise important questions about the relationship between speech perception and phonology –particularly with respect to how perceptually distinct elements on a salience hierarchy need to be in order to be useful to a phonological system– and strongly underline the importance of independent experimental work to this growing area of phonology.

Background: Jun (1995, 2004)

Jun notes a cross-linguistically common pattern of place assimilation in which a coronal stop will assimilate to a non-coronal stop when it is the first consonant in a $V_1C_1C_2V_2$ cluster. In a survey of a number of languages including Catalan, English, German, Hindi, Japanese, Korean, Lithuanian, Malayalam, Yoruba, Zoque and others, the opposite pattern (a non-coronal C_1 assimilating to a coronal C_2) is unattested. Further, among non-coronal stops, /p/ is more likely to undergo assimilation than /k/.

Example: Place assimilation in Korean (cf. Jun 1995, Hume et al. 1999):

a.	/mit + ko/	[mikk'o] "believe and"
	/mit^h + pota/	[mipp'ota] "more than the bottom"
b.	/ip + ko/	[ikk'o] "wear and"
	/nop + ta/	[nopt'a] *[nott'a] "high"
c.	/nok + ta/	[nokt'a] *[nott'a] "melt"
	/kuk + pota/	[kukp'ota] *[kupp'ota] "more than soup"

Jun argues that "…the patterns of place assimilation represent the grammatical reflexes of the requirements of articulation and perception…the crucial basis of the explanation lies in asymmetries in the perceptibility of place of articulation in different segments…" (Jun 2004; 2, 22).

Jun proposes that the pattern is reflective of phonetic knowledge shared by all speakers and listeners regarding the relative perceptual salience of unreleased final stops; in particular, dorsals are more perceptible than labials, and labials are more perceptible than coronals (dorsal > labial > coronal).

In the spirit of the Production Hypothesis (cf. Kohler 1990), this proposal also assumes that articulations with less perceptual payoff will be less resistant to reductive processes such as deletion and assimilation – i.e., speakers are less willing to expend articulatory effort on sounds not easily perceived by listeners.

It is shown that such patterns can be captured elegantly in Optimality Theory by assuming faithfulness constraints which require place features to appear in the output. "PRESERVE", however, is ranked individually for each place, and is in conflict with an energy-conservation constraint, intended to reduce such articulations:

PRES(place(x)): preserve features for place x.

WEAKENING: conserve energy; do not pronounce some or all features of segment x.

Finally, according to Jun (1995), there is a single, universal ranking of the Preservation constraints; all observed patterns of regressive place assimilation in VC_1C_2V sequences in various languages are derived by allowing individual languages to rank WEAKENING differently:

This project was motivated by a set of experiments carried out with Beverley Goodman at Eastern Michigan University. I am grateful for discussions with her, as well as with Jaye Padgett. I would also like to thank those who participated in this study, and the Max Planck Institute for Cognitive and Brain Sciences for use of lab space. Comments and questions are especially welcome and should be sent to **jasonbishop06@fulbrightweb.org**.

1.) Dorsals, labials and coronals all assimilate (Mandarin (Kochetov 2006); Malay, Yoruba and Thai):

WEAKENING **>> PRES(pl(dorsal) >> PRES(pl(labial) >> PRES(pl(coronal)**

2.) Labials and coronals -but not dorsals- assimilate (Korean and possibly some Inuktitut dialects (Dorais 1986)):

PRES(pl(dorsal) >> ***WEAKENING*** **>> PRES(pl(labial) >> PRES(pl(coronal)**

3.) Coronals assimilate, but not labials and dorsals (English, Catalan, German and Yakut):

PRES(pl(dorsal) >> PRES(pl(labial) >> ***WEAKENING*** **>> PRES(pl(coronal)**

4.) No assimilation (Russian (Kochetov 2006)):

PRES(pl(dorsal) >> PRES(pl(labial) >> PRES(pl(coronal) >> ***WEAKENING***

Evidence for Relative Perceptibility

This set ranking of Preservation constraints in combination with a conflicting Weakening constraint is able to generate the correct typological patterns. The source of the set ranking is said to lie in universal phonetic knowledge about the relative perceptibility of unreleased voiceless stops. A VC transition to a coronal is characterized by rapid tongue movement, while labials and dorsal VC transitions are slower, more sluggish gestures with richer acoustic information (dorsals additionally have F2 – F3 convergence which could serve as a cue). The salience hierarchy was deduced primarily from such acoustic and articulatory facts about unreleased voiceless stop consonants; the most satisfying evidence for perceptibility, however, will come from independent perceptual experiments.

Why is this important?

As far as the ranking of constraints is rooted in phonetics, and is reflective of the speaker's sensitivity to the needs of communication (cf. Kingston & Diehl 1994) the proposed hierarchy serves as an explanation for the typological pattern of place assimilation. Yet if the scale is not supported by phonetic evidence regarding perception, the ranking has little independent motivation (other than the fact that it works). Independent perceptual evidence is needed to establish the salience hierarchy above as part of phonetic knowledge.

Previous Research on the relative salience of unreleased final voiceless stops.

Findings of previous studies:

Final unreleased /p/ is often found to be most perceptible, followed by final unreleased /t/. Final /k/ is least well perceived, often considerably less than the other stops. This is in direct conflict with Jun's (1995) predictions.

"Though in general the correct scores are better than chance, several are much worse, as if there were some suicidal feature in the stop (generally k) that makes it sound more like either of the others than like itself…" (Householder 1956; 242)

"The suppression of the releases results in a considerable reduction in identification, especially for k, which…ceases altogether to be identified correctly by the majority of listeners." (Malecot 1958; 377)

Problems with previous studies:

1.) They often have not controlled for the effect of V_1.

2.) Studies have gauged accuracy by using a simple percentage of "hits"; no study has controlled for false positives in any way.

More than likely, vowel environment is relevant. As /p,t,k/ have virtually no internal cues (stop closure is heard as silence), VC transitions are the sole source of information for recovering place contrasts when the stop lacks a release. Indeed the salience of these transitions is the basis for Jun's proposal. Furthermore, a quick comparison of results from some previous studies on the perceptibility of unreleased /p,t,k/ (see next page) suggests that response bias may obscure measures of accuracy which only take into account correct responses (i.e., "hits"). The "calibrated" results from a number of studies below weigh hits against "false positives" (incorrect uses of a given response), and show that /k/ is not least perceptible, but is quite consistently most perceptible:

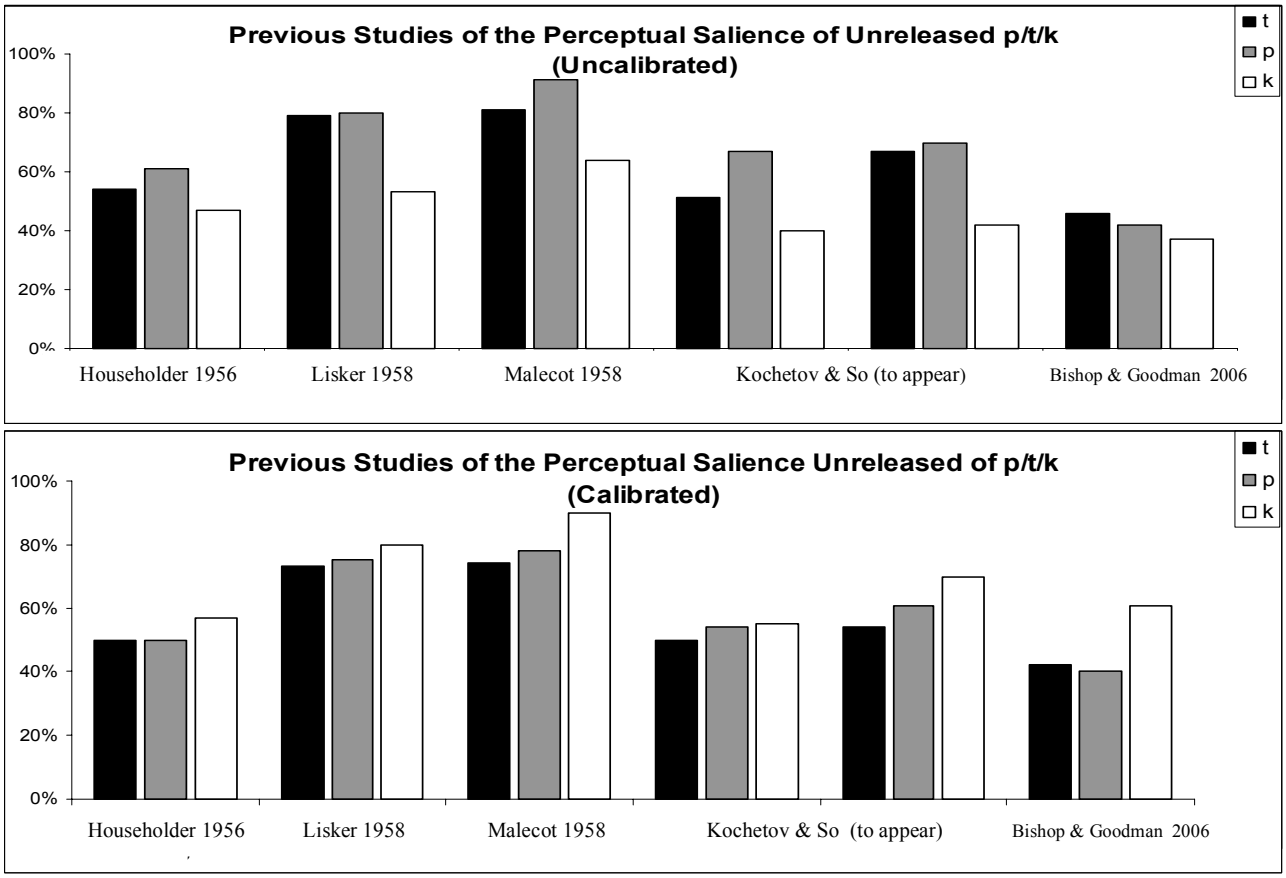

Figure 1. Group results for several studies on the perception of unreleased syllable final /p,t,k/. Lisker (1958) is reported in Lisker (1999); the data from Kochetov & So (to appear) is an average of accuracy for unreleased stops in two different C_2V_2 contexts: before /pV/ and before /kV/. Also, there are two separate experiments shown for that study (experiment 1 on the left, experiment 2 on the right; both test unreleased voiceless stop perception. See Kochetov & So (to appear)). The calibrated versions of the results were estimated using the *group totals* reported in confusion matrices from each study. Thus, because individual participant data were not used in making this overview (and, in most of these cases, such data are unavailable), these figures cannot be subjected to statistical analysis. The comparison here is merely intended as an estimation. The method of calibration is described below under "results".

Present Perception Experiment

A new perception experiment was carried out in an effort to fill in the methodological gaps mentioned in previous work on unreleased final /p,t,k/.

Participants

Participants were 13 speakers of British English, all university students.

Materials & Methods

Recordings were taken in a sound-attenuated booth of a speaker of British English pronouncing nonsense syllables which ended in either /p,t,k/. The stops were placed in seven vowel environments (5 monophthongs and 2 diphthongs; see below). These syllables were then edited using Praat; releases were truncated. The result was seven minimal triplets, differing only in a final voiceless stop, for which the release burst had been removed. Following a trial period, each participant heard six pseudorandom repetitions of the stimuli, resulting in 1,638 judgments (3 stops x 7 vowel environments x 6 repetitions x 13 participants = 1,638). Subjects were to respond in a forced choice fashion after hearing each item.

Stimuli:							
	[at]	[ot]	[ut]	[ɛt]	[it]	[aɪt]	[eɪt]
	[ap]	[op]	[up]	[ɛp]	[ip]	[aɪp]	[eɪp]
	[ak]	[ok]	[uk]	[ɛk]	[ik]	[aɪk]	[eɪk]

Results:

Accuracy was calculated for /p,t,k/ responses for each participant by weighing the ratio of hits against the total number of instances of the response, including "false positives". In doing so, the following formula was used (cf. Grier & Brown 1972; Hume el al. 1999):

Accuracy = (1 – P(fp) + P(hit))/2
Where P(hit) = P('r' | R),
P(fp) = P('r' | X);
'r' is a response,
R is a stimulus having place of articulation 'r', and
X is a stimulus *not* having place 'r'

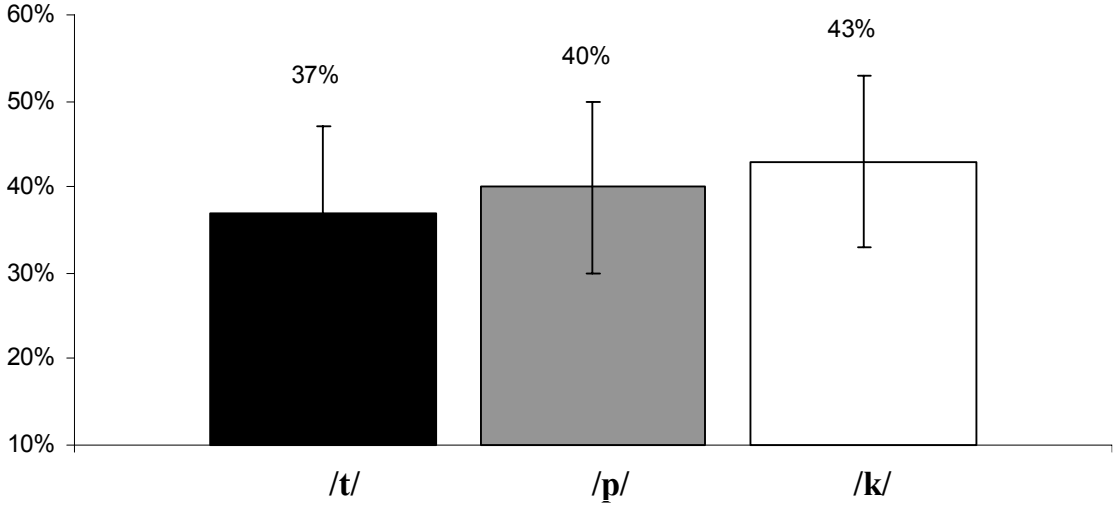

Figure 2. Group averages for accuracy, as defined above, for unreleased final stops across seven vowel environments.

1.) Average accuracy was above chance level (but still quite poor) for each unreleased stop.

2.) Although numerically the perceptibility hierarchy proposed by Jun (1995, 2004) was confirmed in the group average, differences across place of articulation did not reach statistical significance (Repeated Measures AVOVA; ($F(2,12) = 1.392$, $P > .1$)).

Conclusions

1.) The results of previous perceptual studies of unreleased final /p,t,k/ likely reflect:
   a.) effects from preceding vowel place
   b.) a relatively strong response bias (disfavoring /k/)

2.) Controlling for (a) and (b), this set of data suggests that it is unlikely that any differences in the relative perceptibility of final unreleased /p,t,k/ are of the magnitude that would be useful to a phonology.

Further Questions

1.) What is responsible for the response bias?
2.) The interest here has been in perception's influence on phonology; are there phonological facts that may influence perception of final unreleased stops in English?
3.) Just how great a phonetic difference is necessary in order to be useful to a grammar?

References

Bishop & Goodman (2006). The production and perception of American English /p t k/. Paper presented at the Linguistic Society of America Annual Meeting. Albuquerque, NM, 7 January 2006.
Grier, J. (1971). Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological Bulletin*, 75, 424 – 429.
Householder, F. (1956). Unreleased ptk in American English. In Halle, Lunt, McLean, & van Schooneveld: *For Roman Jakobson* (pp. 235 – 244). The Hague: Mouton.
Hume, E., Johnson, K., Seo, M., & Tserdanelis, G. (1999). A Cross-linguistic study of stop place perception. *Proceedings of the XIVth International Congress of Phonetic Sciences,* 2069-2072.
Jun, J. (1995). Perceptual and articulatory factors in place assimilation: an Optimality Theoretic approach. Ph.D. dissertation, UCLA.
Jun, J. (2004). Place assimilation. In Hayes, Kirchner & Steriade (Eds.) *Phonetically-Based Phonology* (pp. 58 – 86). Cambridge: Cambridge University Press.
Kingston, J. & Diehl, R. (1994). Phonetic knowledge. *Language*, 70, 419 – 454.
Kochetov, A. & So, C. (to appear). Place assimilation and phonetic grounding: A cross-linguistic perceptual study. *Phonology,* 24(3).
Kohler, K. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In Hardcastle & Marchal (Eds.) *Speech production and speech modeling*. (pp. 69 – 92) Netherlands: Kluwer.
Lisker, L. (1999). Perceiving final voiceless stops without release: effects of preceding monophthongs versus nonmonophthongs. *Phonetica*, 56, 44 – 55.
Malecot, A. (1958). The role of releases in the identification of released final stops: a series of tape-cutting experiments. *Language,* 34(3), 370 – 380.