Perception and English t-glottalization*

Jeremy O'Brien
jeremypobrien@gmail.com
http://people.ucsc.edu/~jpobrien/

April 13, 2010

1 Introduction

T-glottalization is a phonological process in many varieties of English that realizes /t/ as a glottal stop in certain positions. In this study, a perceptual experiment is used to explain why t-glottalization targets /t/ to the exclusion of other English stops, and why it targets the phonological positions that it does.

There appears to be a large amount of variation when it comes to English t-glottalization. The use of this process can vary by geographical region, social factors, and by phonological position—moreover, the phonological positions vary depending on these other factors. Ladefoged 2006 states that one common position is before /n/ within the word, as in *beaten* [biʔn]. Many varieties of English, including American and British varieties, have glottal stop in coda position in general. For instance, Roberts 2006 provides data on t-glottalization in Vermont English. She finds that the phenomenon is found in speakers of all ages, and that it often occurs at the ends of words, especially pre-pausally. Some varieties allow glottal stop to appear intervocalically as well (which is the case for Cockney English). Docherty & Foulkes 2005 look at glottalization in the English of Tyneside (in northern England). They find that t-glottalization is found in many environments, but only rarely in intervocalic position or pre-pausally. Most of the research on t-glottalization focuses on social factors (age, gender, class, and register distinctions), while there is surprisingly little research about the process in general and the phonological positions it targets.

In the last decade or so, there has been increasing interest in explaining phonological patterns using perception. In particular, many phonologists are exploring the idea of how perceptual phonetics can influence the synchronic grammar. T-glottalization appears to be a perfect candidate for such an analysis. There is no generally accepted abstract/cognitive explanation available (as in features or abstract markedness), and on the surface it appears that [t] and [ʔ] are perceptually similar, so a perceptual explanation seems plausible.

*I would like to express my thanks to my qualifying paper committee: Grant McGuire, Jaye Padgett, and Armin Mester. Also, many thanks to Judith Aissen and the UCSC Winter 2009 Research Seminar for their enthusiasm, support, and comments: Scott AnderBois, Ryan Bennett, Judith Fiedler, Robert Henderson, Matthew Tucker.

This paper presents a perceptual experiment that attempts to explain this connection between [t] and [ʔ], in English and in general. The experiment, a speeded AX discrimination task, explores the perceptual confusability of [t] and [ʔ], and compares it to the confusability between [p] and [ʔ], and [k] and [ʔ] (the other English stop consonants). While *a priori* there are many possible reasons for the development of this alternation (cognitive, articulatory, accidents of history, etc.), the present experiment is designed to test whether perceptual factors play a role.

In investigating a possible perceptual account of the phenomenon, a number of questions must be addressed. The first question is whether perception can serve as a possible motivating factor for t-glottalization. The second question is why /t/ is the target of the process, being replaced with glottal stop in many varieties of English, while /p/ and /k/ are rarely the target of such an alternation. Finally, if a perceptual explanation is possible for t-glottalization in general, then is there also an explanation for the phonological contexts that the process is active in (in particular, focusing more on coda-position than onset-position)?

Anticipating the results, the present investigation found that [t] and [ʔ] are indeed more confusable than other stops compared to [ʔ], but only in coda unreleased position. In this position, the confusability of [k]-[ʔ] was in between that of [t]-[ʔ] (most confusable) and [p]-[ʔ] (least confusable), which suggests that a process of p-glottalization is less likely than k- or t-glottalization. Typologically speaking, no p-glottalization process is attested, but there is a case of k-glottalization found in Indonesian. This process affects /k/ in coda unreleased position, dovetailing nicely with the analysis of perceptual confusability given here.

The terminology for t-glottalization is not necessarily uniform in the literature and in textbooks. The particular type of t-glottalization under investigation here is one where the only point of closure is the glottis. This could be referred to as the glottal replacement or glottaling variety of t-glottalization. The term t-glottalization can also refer to a glottal stop slightly before or concurrent with the alveolar closure. This is called glottal reinforcement or pre-glottalization, and while it is active in many varieties of English, is not the subject of the present study, for the most part because it also effects /p/ and /k/ in many varieties. The glottal replacement form of t-glottalization rarely effects these other stop consonant phonemes.[1] It appears that those varieties of English that have glottal replacement also have glottal reinforcement, and the two choices are varied by phonological and sociological context—this holds for Vermont English and Tyneside English, as cited above. Throughout this paper, the term t-glottalization will be used only in the sense of glottal replacement.

2 Hypotheses and basics of experiment

The present experiment is designed to answer the three questions laid out in the introduction. To decide if perception plays a motivating role in t-glottalization, various forms that contain [t] are compared in the discrimination task to forms that contain [ʔ]. If [ʔ t] is easily confusable, as compared to [ʔ p] and [ʔ k], then this would lend support to a hypothesis that place of articulation of the stop plays a part in the perceptual similarity. This hypothesis,

[1] Brown 1991 states that some varieties will occasionally reduce all the stops /p, t, k/ to glottal stop. My intuition is that, even in these varieties, /t/ still more frequently becomes glottal stop than /p, k/, but it deserves further investigation.

which is a possible answer to the first two questions posed above, is stated in (1).

(1) **Coronal Hypothesis:**
[ʔ t] are perceptually similar, while [ʔ p] and [ʔ k] are more distinct.

Due to the nature of perceptual confusability, we are using the terms *confusability* and *similarity* in a nearly identical sense. Confusability can be an asymmetrical property (e.g. A can be confused for B more often than B is for A), while similarity is usually assessed as a symmetrical property (as in distance along various perceptual dimensions). In the discrimination task explained below, the responses involve only symmetrical cases of confusability, in the form of reaction time values for same/different discriminations.[2] As such, the two terms are more or less interchangeable in this paper.

The hypothesis in (1), along with the symmetrical view of perceptual confusability afforded us by the AX discrimination paradigm, predicts a perceptual map like the one below.

(2) Predicted Perceptual "Map"

[ʔ]	[t]	
[ʔ]		[p]
[ʔ]		[k]

The perceptual experiment was designed to provide quantitative evidence for a map like (2). The experiment is a speeded AX (same/different) discrimination task. The motivation for doing this type of task, as opposed to an identification or rating task, is to access the low-level perceptual system of the participants. In particular, according to the goals of the investigation, the decision of whether the sounds are the same or different should not be mediated by the participants' high-level phonology or cognitive storage systems. The ideal is to see if [t] and [ʔ] are perceptually similar cross-linguistically, with as little influence as possible from categorical perception. Gerrits & Schouten 2004 give a useful discussion on various types of phonetic perception experiments, and the unpredictable way these experiment types access different levels of the perceptual/linguistic system. They talk only briefly about AX experiments. AX experiments are good, they say, because they reduce "the load on auditory memory" by avoiding labeling (as in an identification task). The downside of such an experiment, in their view, is that some participants are heavily biased to "respond 'different' only if they are very sure of their decision", possibly using phonemes as the basis for their decisions. It is arguable that the simplicity of the AX discrimination task, in this situation, outweighs the possible downsides that Gerrits & Schouten 2004 mention. Moreover, the AX task is speeded, where the participants are encouraged to respond quickly and are berated if they respond too slowly. Speeded tasks are sometimes able to bypass language-particular aspects of perception—see Babel & Johnson 2007 and Padgett & Zygis 2007 for support of this notion.

The stimuli (to be expanded momentarily) are of the form [aCa], where C is one of the four stops under consideration: [aʔa, apa, ata, aka].

[2] An asymmetrical metric of confusability might involve the number of times a participant hears [t] and perceives [ʔ], which can differ from the number of times the same participant hears [ʔ] and perceives [t].

The third question under investigation involves the phonological context of t-glottalization. Specifically, many varieties of English have t-glottalization where /t/ would be a coda consonant, while no variety, to the best of my knowledge, has onset /t/ becoming [ʔ] but with coda /t/ being realized as [t]. In order to test the possible perceptual reasons behind this distribution, the experiment compares the intervocalic stimuli from above with coda stimuli. So that the result is more nuanced, and for a basis of comparison, two different types of codas stops are used—released and unreleased. The difference between these two coda types is also important because many, but not all, instances of English coda /t/ are unreleased, either as an unreleased glottal stop (when t-glottalization has taken place) or as an unreleased alveolar [t] (when the rule hasn't taken effect). Taking all of this into account, the final stimuli are given in (3).

(3) Stimuli used in the perception experiment

	ʔ	p	t	k
Intervocalic	aʔa	apa	ata	aka
Coda released	aʔ˹	ap˹	at˹	ak˹
Coda unreleased	aʔ˺	ap˺	at˺	ak˺

The use of these stimuli suggests two competing hypotheses, the Coda Hypothesis and the Release Hypothesis.

(4) **Coda Hypothesis:**
[aʔ˹at˹] and [aʔ˺at˺] are perceptually similar, and [aʔa ata] are perceptually distinct.

(5) **Release Hypothesis:**
[aʔ˺at˺] is perceptually close, and [aʔa ata], [aʔ˹at˹] are perceptually distinct.

3 Methodology

3.1 Stimuli

The stimuli were produced by a trained phonetician and native speaker of English. The utterances were recorded onto a computer at 44.1kHz using Audacity in a sound-attenuating booth. Many tokens of each stimulus type (word) were recorded, and the seven best tokens of each word were chosen as final stimuli in the experiment. Best, in this sense, means free of any sounds uncharacteristic of low vowels and stops, and relatively free of any suprasegmental inconsistency.

The sounds were then amplitude-normalized in Praat, in order to promote greater consistency between the stimuli. This was performed using Praat's "Scale" command, a form of peak amplitude normalization. To this same goal, the files were also duration normalized, such that the duration of stimuli within each condition (intervocalic, released coda, and unreleased coda) was the same. The normalization was done by Audacity, using the "Change Tempo" function. Duration normalization allowed for the reaction times within each condition to be compared. Unfortunately, the sounds could not be compared across conditions, because each condition was assigned a different duration. Even if the durations were normalized to be the same, the various cues to place of articulation would still appear at different times across the three conditions (e.g. the cues for the intervocalic stimuli might come sooner than the cues for unreleased coda stimuli, because there still needs to be time for the following vowel in the intervocalic case).

The final stimuli used are given in (3) above. Representative spectrograms for each of the 12 words are provided in Figure 1. An acoustic analysis of the stimuli is included in Section 4.

3.2 Perception

As shown in (3), there are 12 words in total. These words were broken up into three blocks, according to the post-consonantal condition: intervocalic, released coda, and unreleased coda. For the reasons explained above, comparisons were only made within these blocks. In other words, there were no comparisons between [apa] and [aʔ˺], or between [at˺] and [ak˺]. The order of the blocks was manipulated so that each of the six orderings was used at least once. The following charts give all the possible combinations for each of the three blocks.

(6) Charts of possible stimulus combinations

Block One: Intervocalic

aʔa	aʔa	apa	aʔa	ata	aʔa	aka	aʔa
aʔa	apa	apa	apa	ata	apa	aka	apa
aʔa	ata	apa	ata	ata	ata	aka	ata
aʔa	aka	apa	aka	ata	aka	aka	aka

Block Two: Released

aʔ⌜	aʔ⌜	ap⌜	aʔ⌜	at⌜	aʔ⌜	ak⌜	aʔ⌜
aʔ⌜	ap⌜	ap⌜	ap⌜	at⌜	ap⌜	ak⌜	ap⌜
aʔ⌜	at⌜	ap⌜	at⌜	at⌜	at⌜	ak⌜	at⌜
aʔ⌜	ak⌜	ap⌜	ak⌜	at⌜	ak⌜	ak⌜	ak⌜

Block Three: Unreleased

aʔ˺	aʔ˺	ap˺	aʔ˺	at˺	aʔ˺	ak˺	aʔ˺
aʔ˺	ap˺	ap˺	ap˺	at˺	ap˺	ak˺	ap˺
aʔ˺	at˺	ap˺	at˺	at˺	at˺	ak˺	at˺
aʔ˺	ak˺	ap˺	ak˺	at˺	ak˺	ak˺	ak˺

Within each block, there are 16 possible AX stimuli pairs. The various pairs were repeated many times to the participant, with the "same" pairs weighted so that same and different pairs occurred the same amount of time (84 pairs for "same", 84 pairs for "different", within each block). There was a 100 ms interstimulus interval between the two words of the stimuli. Reaction time was measured from the beginning of the second word. This means that participants could answer before the second sound had finished playing, but after the relevant cues were interpreted so that a decision could be made.

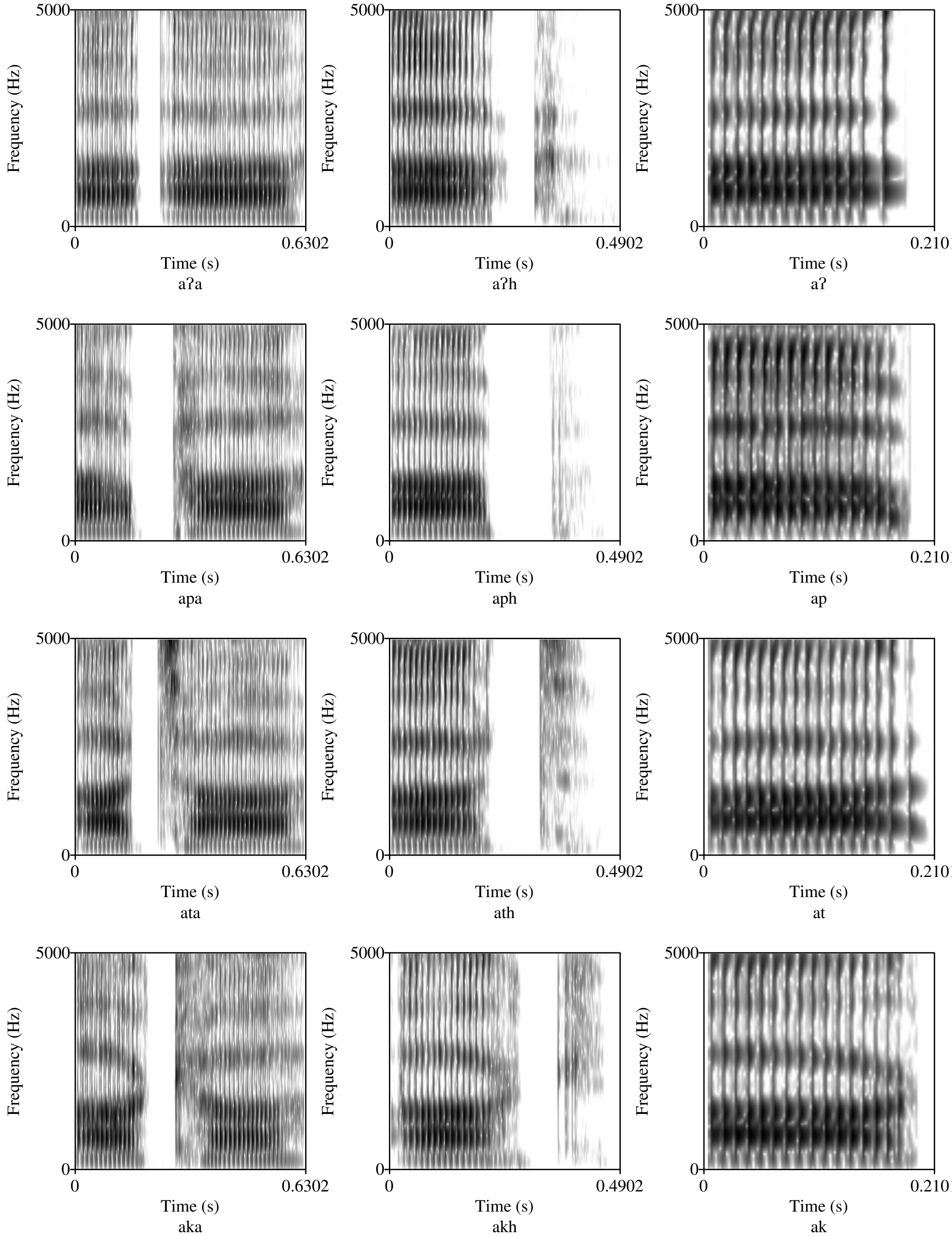

Figure 1: Representative spectrograms for each stimulus word

3.3 Participants

There were 11 listeners in this study. Six were female, and five were male, and all of them were students at the University of California, Santa Cruz. Participants were compensated for their time, receiving either course credit or $5. Of the participants, three were native Spanish speakers, seven were native English speakers, and one was native in both English and Armenian. All participants were fluent or mostly fluent in English. The results for the native Spanish speakers did not appear drastically different than the native English speakers (see Figure 5 below). This in part supports the use of speeded AX discrimination tasks to collect language-independent perception data.

The perception experiment was conducted using SuperLab 4.0 (Mac version) in a sound-attenuating booth. With directions and short breaks between blocks, the experiment took approximately 35 minutes for each participant.

4 Results

The listeners performed very well in responding correctly to the same/different discrimination task. Overall, 95% of the "different" tokens were correctly categorized as different by the listeners.[3] As such, percentage errors are not shown or examined. Instead, reaction time is used to indicate perceptual similarity. Long reaction times on correct comparisons are taken to mean perceptual similarity, while short reaction times are interpreted as the two sounds being perceptually distinct.

Data was pre-processed in the following manner. Only correct "different" responses for the three relevant comparisons were analyzed. That is, only items that compare [p ʔ], [t ʔ], or [k ʔ] and were correctly identified as different sounds were included in the data analysis. Those results whose reaction times were greater than two standard deviations above the mean were also excluded, on the assumption that they were outliers caused by distraction, a long period of contemplation about the sound comparison, or other such factors that are undesirable in a speeded discrimination task. There were 30 of these extra-long reaction times, accounting for only 2.2% of the correct relevant responses. By the measure of being two standard deviations from the mean, there were no outliers that were too short in reaction time.

After removing outliers, the reaction times approximated a normal distribution. Mean reaction time for relevant correct "different" responses was 748 ms (sd = 176ms).

The three graphs below, Figures 2 – 4, show the mean reaction times of the relevant comparisons, where each figure corresponds to an experimental condition. The error bars represent a confidence interval of 95%. A two-way ANOVA was performed, with reaction time as the dependent variable and experimental condition and sound comparison as the two independent variables. An effect of condition and sound comparison were found (condition: $p < 0.01$, $F = 422.3$, $df = 2$; sound: $p = 0.080$, $F = 1.97$, $df = 5$). More importantly for the present purposes, a combined effect of condition by sound was also found ($p < 0.01$, $F = 3.25$, $df = 10$). This provided sufficient evidence to look at these conditions individually

[3]For "same" tokens, 89% were correctly identified as being the same sound. The overall rate of successful identification for the entire experiment was 92%.

to see how the sound comparison and condition interact.

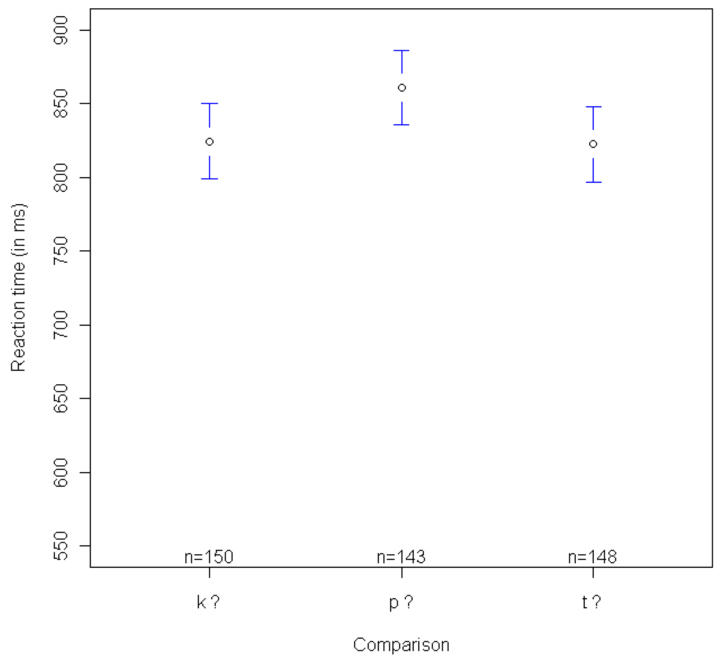

Figure 2: Reaction time for three relevant comparisons, condition VCV

The results were divided into three subsets, each corresponding to a different experimental condition (as the three included figures show). A one-way ANOVA was run on each subset. For the VCV condition, no effect of sound comparison was found ($p = 0.065$, $F(2, 438) = 2.75$). On the other hand, an effect was found for the other two conditions (VC$^{\ulcorner}$: $p = 0.003$, $F(2, 431) = 5.73$; VC$^{\urcorner}$: $p = 0.042$, $F(2, 415) = 3.19$). For each of the two coda experimental conditions, a Tukey honestly significant difference test was performed. In the released condition (Figure 3), the comparison [t ʔ] was significantly lower in reaction time than the comparisons [p ʔ] and [k ʔ]. In the unreleased condition (Figure 4), the only significant difference of means was that between [t ʔ] and [p ʔ], where the reaction time for [t ʔ] was longer.

The results for the three native Spanish speakers were more or less in line with the native English speaker participants. In Figure 5, the mean reaction times for participants 15, 19, and 20 are compared with the means across all the other participants. The left column shows the intervocalic condition, the middle column shows the released coda condition, and the right column shows unreleased. The lowest row are the pooled results for all native English speakers in the study.

When the native Spanish speakers were removed from the data analysis, the results did not change drastically. Some of the p-values increased, but none of the previously significant p-values increased above 0.1. (Some increase in p-values is to be expected because the sample size decreased by 27%.)

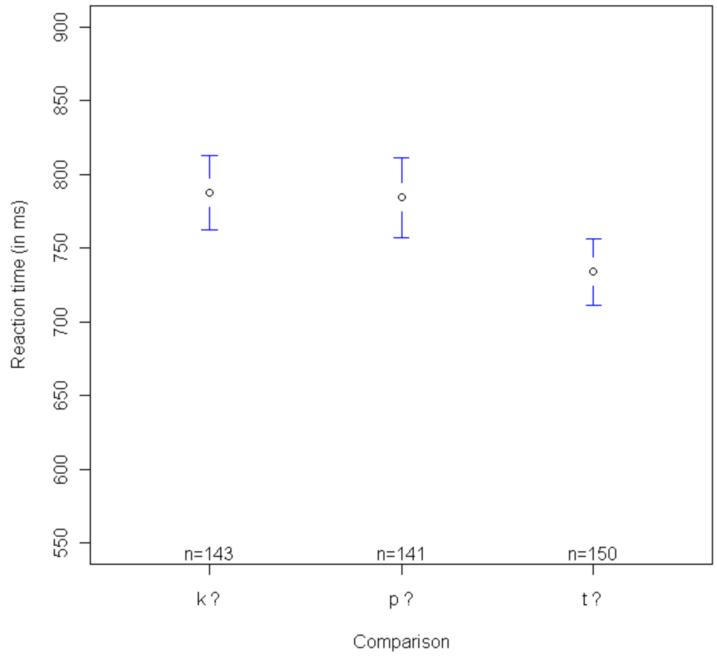

Figure 3: Reaction time for three relevant comparisons, condition VC⌜

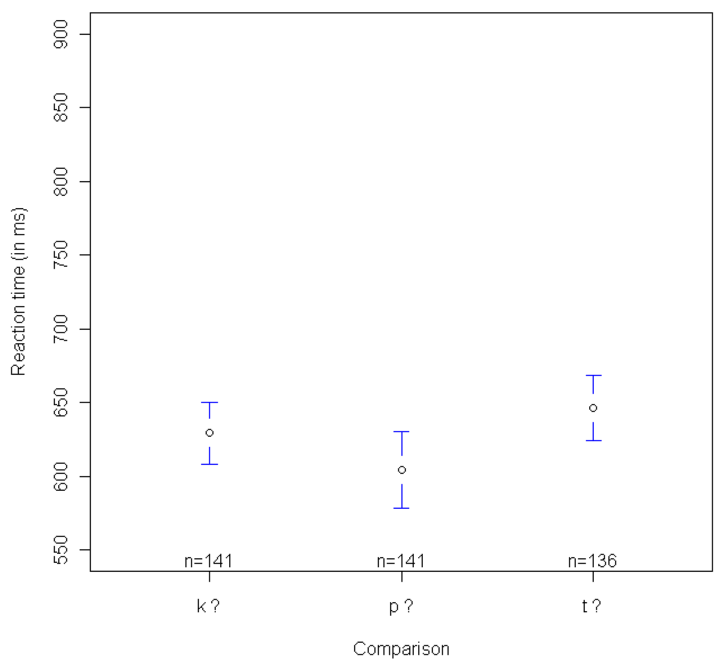

Figure 4: Reaction time for three relevant comparisons, condition VC⌝

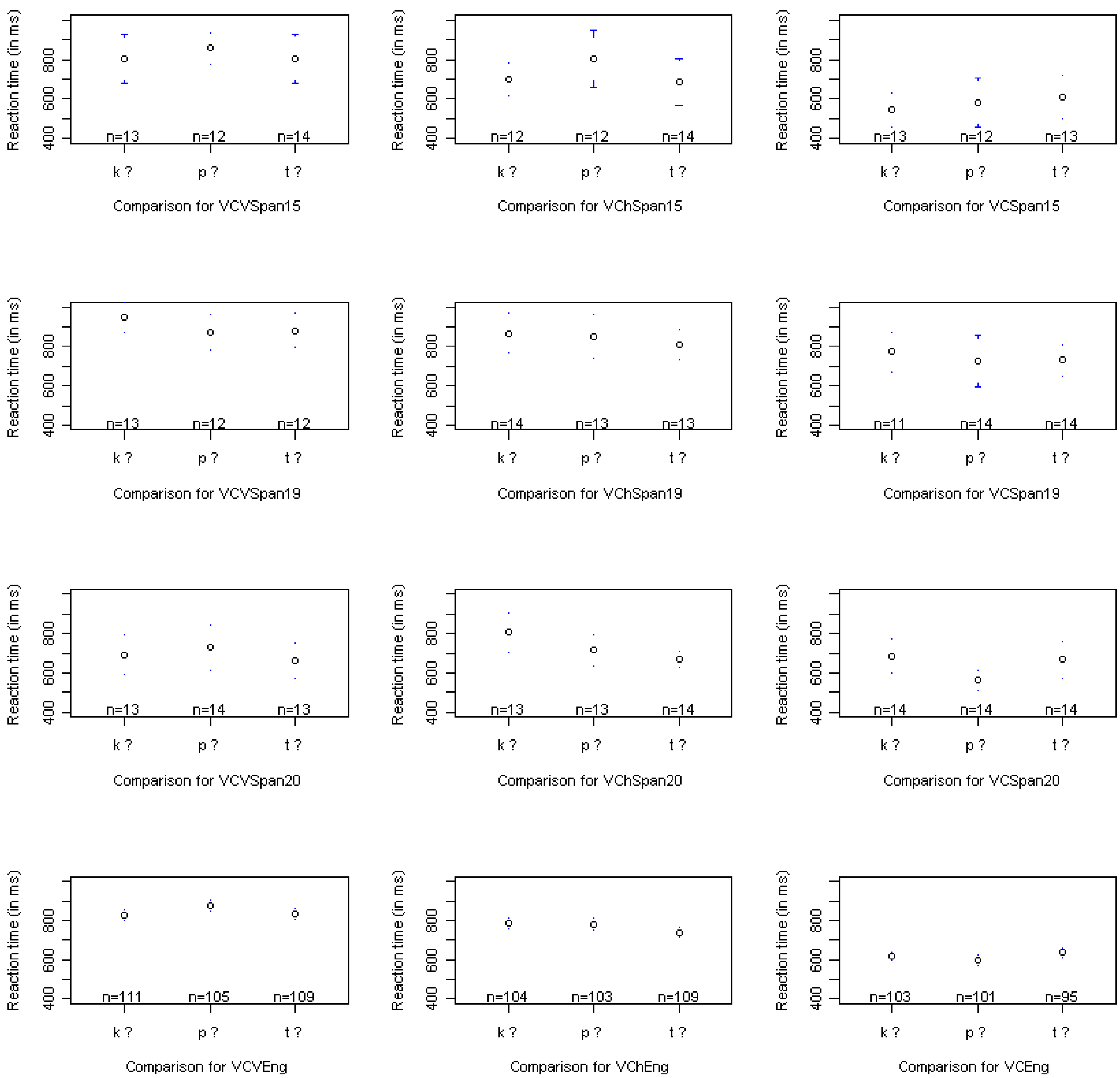

Figure 5: Comparisons between the three native Spanish speakers and the remaining native English speakers

5 Discussion

The results of the analysis demonstrate no statistically significant effect of place of articulation on the confusability of intervocalic [p t k] to [ʔ]—it appears that all three English phonemic stops have about the same similarity to glottal stop when between vowels. When in a released coda position, [p] and [k] are more similar to [ʔ] than [t] is to [ʔ]. This is the exact opposite claim of the Coronal Hypothesis in (1). When looking at unreleased codas, [t] and [ʔ] are more perceptually similar than [p] and [ʔ]. This provides partial support for the Coronal Hypothesis.

In terms of the Coda and Release Hypotheses (4) – (5), it appears that the Release Hypothesis has some support from the present experiment. While [t] and [ʔ] are perceptually distinct when released, they are similar when they are unreleased (at least when compared to [p] and [ʔ]). On the other hand, there is no support at all for the Coda Hypothesis.

While the experiment was only able to find partial support for the Coronal Hypothesis, this is not necessarily a surprising result. There are a lot of cues available to the listener when comparing released consonants, cues that are not available for unreleased ones. In particular, listeners can use the spectral properties of the release burst, and the intensity of that burst, to determine place of articulation. Furthermore, the CV formant transitions are available to the intervocalic consonants. The only cue to place of articulation that is afforded the unreleased consonants is the VC formant transition, which is also there for the other experimental conditions. One explanation that is consistent with the present results is that [t] and [ʔ] are perceptually distinct when burst-related cues can be used, but when those cues are absent, they are perceptually similar. The spectrograms in Figure 1 support this explanation. The release burst of [t] in [ata] and [at˹] shows a large amount of high-frequency noise, in the area above 4 kHz. This sibilant-like cue for place of articulation in stops is not found in any of the other stops. The burst for [p] is quite weak, and the bursts for [k] and [ʔ] show a fairly uniform spectrum of noise. The sibilant nature of released [t] might be what keeps it from being similar to released [ʔ].

At the same time, an acoustic analysis of the stimuli brings up a complicating factor. The unreleased [at˺] in Figure 1 shows a small amount of creaky voice, where the final glottal pulse appears after a small lag. This feature is not found on any of the tokens for [ap˺] or [ak˺], but it is shared with tokens of [aʔ˺]. Of the seven [at˺] tokens used in the experiment, four of them show this creaky voice. This provides an alternative hypothesis, whereby the participants are using this creaky voice cue to evaluate the sounds, and in so doing, find the creaky voiced [at˺] tokens to be similar to glottal stop. There is very little within the experiment that can tell us what cues the listeners are using, so this must be left an open question.

6 Comparison with Indonesian

The results of the present experiment suggest that there is a language-universal perceptual reason for English t-glottalization to target /t/ to the exclusion of /p/ and /k/. Specifically, unreleased coda [t] and [ʔ] are perceptually confusing, and therefore perceptually similar. However, the only statistically significant comparison from the unreleased coda experimental condition was actually that between [t ʔ] and [p ʔ], where the pair [t ʔ] was more similar. None of the comparisons involving [k ʔ] were statistically significant. This translates into cross-linguistic predictions in the following manner: when the environment involves unreleased stops, t-glottalization is much more motivated than p-glottalization, and therefore it should be a more common process, and those language varieties that have p-glottalization should also have t-glottalization. No prediction can be made about the relative frequency or naturalness of k-glottalization, however.

Indonesian, then, presents an interesting data point, consistent with but not reliant on our cross-linguistic predictions. In Indonesian, it is the voiceless *velar* stop that reduces to [ʔ], to the exclusion of /p/ and /t/. Furthermore, coda stops in Indonesian are always unreleased (Lapoliwa, 1981). The existence of Indonesian k-glottalization, in unreleased coda environments, is consistent with our results because it does not favor p-glottalization over t-glottalization. With further study, there remains the possibility that unreleased [t] and [k] are approximately the same distance from [ʔ]. If that were the case, then perhaps languages can choose which stop becomes [ʔ], or maybe there is another language out there that allows both stops to reduce.

7 Conclusion

It appears that t-glottalization is only one of many phonological processes that target /t/ in English. There are a number of phonological processes that weaken /t/ to the exclusion of other phonemes:

(7) T-glottalization: the process under consideration here

(8) Post-nasal t-deletion: *winter* → [wɪnr]

(9) Flapping: *pretty* → [prɪɾi] (which targets both /t/ and /d/)

The final processes target many segments, but impressionistically they seem to target /t/ quite frequently.

(10) Assimilation to following obstruent: *short cut* → [ʃɔrkkʌʔ]

(11) Fast speech deletion: *listless* → [lɪsləs] (Raymond et al., 2006)

A possible explanation for this distribution, related to Universal Grammar or inherent cognitive mechanisms, is that /t/ is underspecified for most phonological features. If this were the case, then this underspecified consonant would be realized as [t] in some environments, and [ʔ] in others, which makes t-glottalization look like a very normal and unmarked phonological process.

Alternatively, /t/ could be the phoneme that is weakened most often because it is frequent. This line of thinking is explored in Beckman et al. 2003, where patterns of acquisition for English and Japanese are compared. English-learning children will commonly front /k/ to /t/, but the opposite pattern is found for many Japanese-learning children. As such, a language-universal explanation (such as simply stating that /t/ is less marked) is not able to account for the difference. Language-particular characteristics, like phoneme frequency, can account for this difference, however. Beckman et al. 2003 report that, in English child-directed speech, /t/ is more frequent than /k/, while Japanese child-directed speech shows that /k/ is more frequent. These acquisition facts might, then, be related to /t/ being the target of so many weakening processes in English.

The cognitive reasons mentioned above might provide a sufficient answer to why t-glottalization targets /t/. However, articulatory or perceptual effects may also play a role (either on their own, or in conjunction with cognitive factors)—in fact, an explanation grounded in perceptual facts is the one pursued in this paper. One possible explanation of t-glottalization is that perception is the direct motivating force: listeners mishear [t] as [ʔ], and over time they come to produce it this way. The perceptual similarity first causes some tokens of (coda, unreleased) [t] to be perceived as [ʔ], but then this misperception affects the way speakers internalize and then later produce the sound. This type of argument is put forth in Ohala 1981, where some sound changes are shown to arise from over- or under-application of reconstruction rules related to overlapping auditory cues.

An alternative approach would be to say that speakers try to minimize articulatory effort whenever they can get away with it. One way to minimize effort would be to reduce a /t/ to a [ʔ]. What was once an alveolar gesture becomes a "gestureless" stop consonant. Perhaps speakers can get away with a reduction of /t/, because it is perceptually similar to [ʔ], but they can't get away with /p/ or /k/ reducing. This line of reasoning can be found in Liljencrants & Lindblom 1972, where a computational model of articulation is combined with a model of maximizing auditory contrast. Likewise, Kohler 1990 follows a similar argument—German consonant reduction rules are explained as eliminating the need for finely-controlled motor movements, but only if they do not contain a "high signalling value" such as syllable-initial consonants. One reason to be suspicious of this line of reasoning is because coronals are common sounds cross-linguistically, presumably because they are easy to articulate. If this presumption is correct, then it is not obvious that /t/ should be the target for lenition, as opposed to some more difficult-to-articulate consonant. Of course, in a sophisticated model this force would be counteracted by /t/'s frequency in English. (See Winters 2003 for a discussion on the difference between Ohala-style and Liljencrants & Lindblom-style explanations.)

Either style of analysis is compatible with the experiment described above. That is to say, even if the results unambiguously pointed to [t] and [ʔ] being perceptually similar, we still would not be able to decide between these two competing explanations for how perception can effect the grammar. The present experiment does, however, give support for perceptual accounts for this process in general.

Also important is the fact that perceptual similarity usually has some type of underlying reason. That is, two sounds are perceptually confusable for some more fundamental reason. This might be articulatory—for instance, the coronal gesture is generally faster than dorsal or labial gestures, so the duration of the closure could make [t] less distinct from a (mostly gestureless) glottal stop. Alternatively, the confusability of [t] and [ʔ] could involve the formant transitions into and out of the stops, which are a result of articulatory and acoustic realities. Or perhaps some other cause is ultimately to blame for their similarity. Regardless of the underlying source, we can still probe to see if the primary explanation of t-glottalization is perceptual. A more nuanced approach in future experiments could provide some evidence for secondary explanations.

References

Babel, Molly & Keith Johnson (2007) Cross-linguistic differences in the perception of palatalization. *US Berkeley Phonology Lab Annual Report* : 265–268.

Beckman, Mary E., Kiyoko Yoneyama, & Jan Edwards (2003) Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. *Journal of the Phonetic Society of Japan* **7**: 18–28.

Brown, Adam (1991) *Pronunciation Models*. Singapore University Press.

Docherty, GJ & P. Foulkes (2005) Glottal variants of /t/in the Tyneside variety of English. *A figure of speech A festschrift for John Laver* : 173–197.

Gerrits, E. & M.E.H. Schouten (2004) Categorical perception depends on the discrimination task. *Perception & Psychophysics* **66**(3): 363–376.

Kohler, Klaus J. (1990) Segmental reduction in connected speech in German: phonological facts and phonetic explanations. *Speech production and speech modelling* **55**: 69–92.

Ladefoged, Peter (2006) *A course in phonetics, fifth edition*. Thomson Wadsworth.

Lapoliwa, Hans (1981) *A generative approach to the phonology of Bahasa Indonesia*. Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Liljencrants, Johan & Bjorn Lindblom (1972) Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* : 839–862.

Ohala, John J. (1981) The listener as a source of sound change. *Chicago Linguistic Society parasession on language and behavior* : 178–203.

Padgett, Jaye & Marzena Zygis (2007) A perceptual study of Polish fricatives, and its relation to historical sound change. *Manuscript in preparation, University of California, Santa Cruz* .

Raymond, William D., Robin Dautricourt, & Elizabeth Hume (2006) Word-internal/t, d/deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change* **18**(01): 55–97.

Roberts, Julie (2006) As old becomes new: Glottalization in Vermont. *American Speech* **81**(3): 227–249.

Winters, Steven (2003) *Empirical investigations into the perceptual and articulatory origins of cross-linguistic asymmetries in place assimilation*. Ph.D. thesis, The Ohio State University.