The patterns of vowel insertion in IL phonology: The P-map account*

Bo-Young Kwon
(Michigan State University)

Kwon, Bo-Young. 2005. The patterns of vowel insertion in IL phonology: The P-map account. *Studies in phonetics, phonology and morphology* 11.2. 21-49. In this study, I provide evidence that the P-map (Steriade 2001a, 2001b), a model of generic listeners' perceptual abilities and biases, is responsible for the patterns of vowel insertion in Korean speakers' production of English consonant clusters. The experimental findings in the study demonstrate that vowel insertion is more likely to occur a) when C_1 (the first consonant of a cluster) is a stop than a strident or a sonorant and b) when C_1 is a voiced stop than a voiceless stop. This study presents two proposals to account for the attested patterns of vowel insertion. First, following Côté (2000), I propose that the likelihood of vowel insertion in consonant clusters correlates with the perceptual salience of a segment in contrast: the weaker the perceptual cues of a segment, the more likely it is to be modified. In a similar vein, I also propose that second language (L2) learners are guided by the P-map when faced with L2 phonotactic constraints that are not part of their native languages. That is, the P-map knowledge leads L2 learners to choose the output form with relatively minimal modification of the input in terms of perceptual similarity. These two proposals are supported by the author's auditory similarity judgment experiment involving the same population as in the production study. This perception experiment shows that there exist perceived similarity differences between unepenthesized input and epenthesized output and these perceived similarity differences actually correlate with the patterns of vowel insertion in Korean speakers' production data. **(Michigan State University)**

Keywords: L2 acquisition of English consonant clusters, vowel insertion, P-map, auditory similarity judgment

1. Introduction

This study seeks to find principled explanations for interlanguage (hereafter IL) sound modification phenomena, focusing on Korean speakers' strategies for realizing English consonant clusters (e.g., ***gl**ass*, *te**st***) and the related issues of variation in their production. In particular, in this study, I provide evidence that the P-map (Steriade 2001a, 2001b), a model of generic listeners' perceptual abilities and biases, is responsible for the patterns of vowel insertion demonstrated in Korean speakers' production of English consonant clusters.

Difficulties that second language (hereafter L2) learners face in the

* I would like to thank two anonymous reviewers for their valuable comments on this paper. I am also grateful to Prof. Yong-soo Hwang for helping me with recruiting participants for the experiments, to Kim Lewis for data analysis, and to Li-Jen Shih for constructive discussion on the topic. An earlier version of this paper was presented at the Annual Meeting of the Chicago Linguistics Society 41, April 7-9 2005 University of Chicago, IL.

production and perception of second language sounds that are not part of their first language (hereafter L1) are well established. The acquisition of consonant clusters by L2 learners (e.g., Japanese, Korean, and Mandarin Chinese) whose native languages do not allow consonant clusters is a good example for demonstrating this kind of difficulty. Previous studies on the production of consonant clusters by L2 learners have consistently reported that L2 learners insert a vowel or delete a consonant to break down L2 consonant clusters (e.g., Abrahamsson 2003, Anderson 1987, Broselow 1987, Broselow and Finer 1991, Eckman 1981a, 1981b, Eckman and Iverson 1993, Hansen 2001, Kim 2000, Tarone 1987, among others). At first glance, it seems to be the clear case of first language interference in L2 acquisition. However, this phenomenon is more interesting because asymmetries occur in the use of strategies for realizing consonant clusters.

First, there occur more sound modifications in coda clusters than in onset clusters (Hancin-Bhatt and Bhatt 1997: 345, Kim 2000: 207, Yoo 2004: 487). Second, vowel epenthesis is more likely to occur word-initially, whereas deletion is more likely to occur word-finally (Hancin-Bhatt and Bhatt 1997: 345). Third, the less sonorant segment in a cluster is usually deleted (Hansen 2001: 356, Tropf 1987: 185). Fourth, vowel epenthesis is more likely to occur in a cluster of voiced obstruent + C than in a cluster of voiceless obstruent + C (Kim 2000: 138).

Now, questions arise regarding the above observations: **a)** Why are coda clusters more favored targets of sound modification than onset clusters? **b)** Why is an epenthetic vowel frequently inserted in some positions, but not in other positions? **c)** Why are some types of segments in a cluster more likely to be the target of consonant deletion? **d)** Can these patterns be generalized cross-linguistically or is each pattern a mere reflection of L2 learners' first language transfer? **e)** Are these patterns predictable to any degree? It is crucial that these questions be answered, given that previous studies have not been very successful in providing comprehensive and consistent accounts of the phenomena at hand.

In fact, the difficulty of accounting for highly variable L2 learners' production data under the theoretical linguistic frameworks has been a continuing issue in L2 research. For instance, Broselow (1987) points out that the strong form of the Contrastive Analysis Hypothesis (hereafter CAH) (Lado 1957) fails to predict what the language learners would do to resolve the problems encountered in the course of L2 acquisition. That is, CAH may explain some of the L2 phonological errors post-hoc as a result of differences between the native language and the target language, but the form of the errors cannot be predicted with any regularity (Broselow 1987: 292). Hansen (2001: 338) also notes that studies typically focusing on linguistic constraints such as markedness (e.g., Eckman and Iverson 1993, Major and Faudree 1996), L1 transfer (Broselow 1987, Sato 1984, 1989) or universals (Benson 1988) do not fully account for the acquisition of L2 phonology, if the analysis is based on one specific linguistic constraint.

As an alternative to previous analyses, I will test the validity of a recently emerging principle, the P-map (Steriade 2001a, 2001b), as an analytical tool in L2 phonological acquisition.

The P-map is a set of statements about perceived distinctiveness-differences between different contrasts in various contexts. According to the P-map proposal, when faced with a phonotactic constraint, speakers prefer, as a solution, “the least distinctive contrast whose modification resolves the violation” (Steriade 2001b: 14). It is proposed in this study that L2 learners are guided by the P-map. The P-map knowledge leads L2 learners to choose the output form with relatively minimal modification of the input in terms of perceptual similarity, while resolving the phonotactic constraints in L2.

Thus, the primary goal of this study is to test the validity of the P-map principle in L2 acquisition of English consonant clusters. Given that L2 learners are trying to produce the forms as closely as possible to a target language, there seems to be a great possibility that the P-map knowledge works very actively in shaping IL phonology. Fan’s study (2004) on Mandarin Chinese speakers’ acquisition of coda clusters demonstrates a good example that the P-map knowledge holds true in IL phonology. Reanalyzing the data from Eckman 1981a, Hansen 2001, Gui 1985, and Weinberger 1987, Fan deals with issues such as a) the choice of an epenthetic vowel: why is a schwa more likely to be inserted rather than other vowels?, b) stop deletion, and c) postvocalic /r/ deletion. For instance, regarding the choice of an epenthetic vowel, Fan proposes that the preference of a schwa as an epenthetic vowel is related to the phonetic features of a schwa. That is, a schwa is the shortest vowel in terms of duration and lacks invariant articulatory properties. Consequently, schwa insertion in a cluster is perceptually least obtrusive (Fan 2004: 19).

At this point, one may ask why L2 learners attempt to change the English input in the first place. Shouldn’t it be optimal that L2 learners’ output is faithful to input without changing any feature in the input? It is true that the ultimate goal of L2 output (or production) is to produce the English input as it is. However, we should also consider that adult L2 learners are already equipped with their L1 phonotactic constraints and have to deal with the mismatch between their L1 and L2 in the course of L2 acquisition. Taking an example from L2 acquisition of consonant clusters, when L2 learners are in the developmental stage where their production of consonant clusters does not obtain ultimate proficiency, they are supposed to make sound modifications in English input to comply with their L1 phonotactic constraint which bans complex margins in a syllable. This is the point where the P-map knowledge plays its role in the system. Among various possible modifications of the English input (e.g., “bret” can be modified as [bet], [ret], [buret], [biret], etc), the P-map knowledge leads L2 learners to choose the one with relatively minimal modification of the input in terms of perceptual similarity.

The second purpose of this study is to establish generalizations of L2 acquisition of English consonant clusters, focusing on the patterns of vowel insertion. Studies specifically aiming to identify the patterns of vowel insertion are rather sparse in L2 research[1], and the data from my study constitute the database for the analysis. The previous studies on the acquisition of English consonant clusters are also referred to so that it can be shown that the P-map account holds cross-linguistically.

In addition, to provide more direct evidence for the P-map proposal in L2 acquisition of English consonant clusters, this study presents an auditory similarity judgment experiment. It will be shown that there is strong correlation between Korean speakers' auditory similarity judgment of unepenthesized input and epenthesized output and the patterns of vowel insertion in their production.

As a formal analysis of the data, Optimality Theory (Prince and Smolensky 1993, hereafter OT) is adopted. The P-map makes crucial connections with the OT phonological model in such a way that the P-map projects faithfulness constraints and determines their ranking in the OT system. The working mechanism of the P-map in OT will be discussed in Section 2.

This paper is organized as follows. In Section 2, I present theoretical frameworks for data analysis in the study: Optimality Theory and the P-map. In particular, I show how OT and the P-map connect to each other as an analytical tool. Section 3 presents the experimental design and results of Korean speakers' production of English consonant clusters. In Section 4, I explain the patterns of vowel insertion in terms of perceptual salience of modified segments and do a formal analysis under the OT framework. In Section 5, I provide supporting evidence for the P-map proposal in the patterns of vowel insertion by conducting an auditory similarity judgment test.

2. Theoretical frameworks

2.1 OT and L2 acquisition of consonant clusters

Optimality theory (Prince and Smolensky 1993) views human languages as a ranked system of conflicting universal constraints. Each of the constraints in the system requires some aspects of grammatical output forms, and the constraint may be either satisfied or violated by an output form. Constraints are universal, but the rankings of the constraints are language specific, which makes cross-linguistic variation possible. Also, constraints are violable, but the violation must be minimal; violation of a higher-ranked constraint incurs more cost than a lower-ranked constraint. Therefore, the candidate with least costly violation is chosen as an optimal

[1] Many of the previous studies investigated L2 acquisition of consonant clusters in terms of relative consonant cluster markedness and L2 learners' error rate.

output.

When it comes to the constraints involved in the selectional mechanism, OT recognizes two types of constraints: *markedness constraints* and *faithfulness constraints*. Markedness constraints require that output forms meet criteria of structural well-formedness. These well-formedness constraints are built upon the notion of markedness. Unmarked values are cross-linguistically preferred and basic in all grammars, while marked values are cross-linguistically avoided and used by grammars only to create contrast (Kager 1999: 3). For instance, complex onsets and codas, consisting of two or more consonants, are universally more marked than simple onsets and codas. Based on this markedness relation in syllable margins, two markedness constraints are generated: *COMPLEX ^{ons} and *COMPLEX ^{cod}

(1) Markedness constraints (Kager 1999: 97)
  a) *COMPLEX ^{ons} ('Onsets are simple')
  b) *COMPLEX ^{cod} ('Codas are simple')

Along with the markedness constraints, OT recognizes faithfulness constraints. The current dominating view of faithfulness is Correspondence Theory (McCarthy 1995, McCarthy and Prince 1995). Correspondence is a relation between two structures, such as base-reduplicant and input-output, where candidate outputs are subject to evaluation together with the correspondent input (McCarthy 1995: 14). Correspondence constraints militate against divergences of input and output along one dimension.[2] Examples of correspondence constraints regarding the complex margins are as follows.

(2) Faithfulness constraints (Kager 1999: 101-102)
  a) MAX-IO: Input segments must have output correspondents.
  ('No deletion')
  b) DEP-IO: Output segments must have input correspondents.
  ('No epenthesis')

MAX-IO requires that, for every output segment, there is an input segment corresponding to it: deletion of a segment is not allowed. On the other hand, DEP-IO is violated if the output has a segment that lacks a correspondent in the input: insertion of a segment in the output is not allowed.

Note that markedness constraints and faithfulness constraints are inherently conflicting in the sense that, whenever some lexical contrast is being preserved by keeping faithfulness constraints, there will be some cost associated in terms of markedness. Comparison between syllable structures of Korean and English shows how the conflicting relation between markedness and faithfulness works in the OT system.

2 Correspondence constraint: the output equals the input for some property *P* (Kager 1999: 67).

English and Korean present two distinct cases of syllable margins. English allows complex margins in any of the positions, generating V, VC, CV, CVC, CV**CC**, **CC**V, **CC**VC, **CC**V**CC**.[3] Note that in English the markedness constraint *COMPLEX is ranked lower than faithfulness constraints. Therefore, with respect to OT constraints, the greater variety of English syllable structures is made possible at the cost of violating markedness constraint *COMPLEX.

On the contrary, Korean ranks markedness constraint *COMPLEX higher than faithfulness constraints, limiting possible syllable types (syllable types in Korean include V, VC, CV, C(G)VC)[4] and using deletion and epenthesis to resolve complex margins. English loanwords in Korean demonstrate a good example of how epenthesis and deletion are employed because *COMPLEX is ranked higher than MAX-IO or DEP-IO in Korean phonology. English loanwords with complex margins are adapted as either vowel insertion or consonant deletion as in (3).

(3) English loanwords with complex margins

***gr**ass*	[kɨrɛs]	vowel insertion
*ma**rk***	[makɨ]	/r/ deletion

In this section, we have discussed the basic tenets of OT such as the basic properties of constraints, confliction between markedness and faithfulness constraints, and universality and violability of constraints. In what follows, the concepts of the P-map and how the P-map knowledge generates context sensitive faithfulness constraints will be discussed. We will also see how these context sensitive faithfulness constraints play a role in the OT framework.

2.2 P-map

The review of the P-map in this section will be mainly based on the work of Steriade 2001b. The concept of the P-map was originally proposed as an attempt to solve the “Too-Many-Solutions” conundrum (in Steriade’s term) which arises when the system of constraints and rankings predicts too many solutions of a given phonotactic problem. The case of the obstruent devoicing process demonstrates this problem. A phonotactic constraint on obstruent voicing is cross-linguistically well-evidenced.

(4) A phonotactic constraint (Steriade 2001b: 2)

*[+VOICE]/_]: voiced obstruents do not occur at the end of the word.

---

[3] In English, inflectional suffixes at the end of words create complicated codas with up to four consonants in a row.

[4] There has been general agreement that Korean syllable structure is composed of $[(C)(G)V(C)]_{\sigma}$ (where G is a glide). However, there exist several different views on the status of glides in a syllable. Some researchers treat Korean glides as part of the nucleus (e.g., Kim and Kim 1991, Kim 1990, Sohn 1987), while others regard them as part of an onset cluster (e.g., Ahn 1985, Lee 1991, Lee 1993).

The crucial observation on which the P-map proposal is based is that even though there are many possible ways of repairing the phonotactic constraint on word-final voiced obstruent (such as voiced obstruent devoicing, nasalization, deletion, metathesis, or post-voiced obstruent epenthesis), only devoicing occurs as a reaction to *[+VOICE]/_] violation. Then, the question is what selectional mechanism works for favoring the 'devoiced' output instead of alternative options of modifying the input.

Steriade (2001b) finds the answer to this question in the P-map, a set of statements about perceived distinctiveness-differences between different contrasts in various contexts. That is, the optimal output, according to Steriade, would be the one which differs from the input minimally when complying with the phonotactic constraints. Following this reasoning, other alternative modifications of input are avoided because they result in drastic input-output differences.

There are two properties of the P-map that are critical to understanding its mechanism: *positional effects* and *contrast*. *Contrast* is the statement such that "*a* is more perceptible than *b*" means "*a* is more reliably distinguished from a reference term *x* than *b* is distinguished from *x*" (Steriade 2001b: 15). *Positional effect* is the notion that the distinctiveness of a segment is affected by the syntagmatic context. That is, contrast in certain phonological contexts shows more faithful effects than contrast in other contexts (e.g., perceptual salience of voicing contrast may vary depending on the context where it occurs).

Based on the working mechanism of the P-map discussed above, Steriade (2001b) proposes that devoicing is chosen as an optimal repair strategy, since the pair [tæb – tæp] (devoicing) is perceptually more similar than other possible modifications (e.g., nasalization [tæb – tæm], deletion [tæb – tæ], metathesis [tæb – bæt], or post-voiced obstruent epenthesis [tæb – tæbə]). All of the other modifications result in greater input-output dissimilarity than devoicing, and therefore they are systematically avoided (Steriade 2001b: 4).

One may wonder at this point how we know that devoicing is the least modification from the input. Actually, this is an important question to address given that the gist of the P-map proposal is that there exists correlation between perceived similarity differences and phonological process. More specifically, it should be shown that "perceived degree of similarity differences correlates with choices made in phonological systems between alternative options of modifying an input" (Steriade 2001b: 6).

When it comes to the question of what factors determine the relative similarity of different contrasts, Steriade follows an inductive approach to similarity in which if the pair *z-w* causes more confusion than the pair *x-y*, then *z-w* is more similar than *x-y*. Simply put, the more confusable the pair is, the more similar it is.

Returning to the issue of similarity between voiced and devoiced obstruent pairs, there is abundant evidence that voicing contrast is perceived as less distinctive than contrast based on obstruency differences (see Greenberg and Jenkins 1964, van den Broecke 1976, Vitz and Winkler 1972, Walden and Montgomery 1975). For instance, in Greenberg and Jenkins's (1964: 168) study, subjects were asked to list associates to nonsense stimuli like [klæb], and it turned out that the most common responses involved voicing changes. Thus, for [klæb], the most commonly mentioned forms were [klæp] (23/46 responses), while the other potential associate, [klæm], which also differs by exactly one feature from the stimulus, was elicited less frequently (11/46 responses). This result apparently shows that difference in nasality ([klæb]-[klæm]) is more distinctive than difference in obstruent voicing ([klæb]-[klæp]).

The P-map is incorporated into the OT by projecting and ranking correspondence constraints in the system. That is, based on the perceptual similarity between the input and output form, the P-map projects faithfulness constraints and determines its ranking. For instance, if a contrast x-y/_A is more perceptible (or distinctive) than a contrast x-z/_A, then the P-map's effect on the grammar will rank higher the correspondence constraint x-y/_A than correspondence x-z/_A. These ideas are schematized in (5).

(5) P-map effects on the ranking of correspondence conditions (Steriade 2001b: 5)

P-map comparisons	More distinctive contrast [b]-[m] in V_] vs.	Less distinctive contrast [b]-[p] in V_]
Ranking of correspondence constraints	Higher ranked constraint IDENT [±nas]/ V_] >>	Lower ranked constraint IDENT [±voice]/ V_]

Table (5) illustrates that IDENT [±nas]/ V_] (the case of final obstruent nasalization) ranks higher than IDENT [±voice]/ V_]) (the case of devoicing) based on the observation that obstruency difference in word-final position is more perceptible than voicing difference in the same position.

Tableau (6) shows how markedness constraint *[+voice]/_] and faithfulness constraints IDENT [±nas]/ V_] and IDENT [±voice]/ V_] interact with each other, resulting in [tæp] as an optimal output.

(6) Final obstruent devoicing

/tæb/	*[+voice]/_]	IDENT [±nas]/ V_]	IDENT [±voice]/ V_]
a. tæb	*!		
☞ b. tæp			*
c. tæm		*!	

In (6), the most faithful output [tæb] loses its candidacy as an optimal output, since it violates the highest ranked markedness constraint, *[+voice]/_]. Between the remaining [tæp] and [tæm], the output [tæp] wins over [tæm] due to the constraint ranking IDENT [±nas]/ V_] >> IDENT [±voice]/ V_], which demonstrates the observation that [tæp] is less deformation of the input /tæb/ than [tæm] is of /tæb/ in terms of perceptual similarity.

In this section, we have discussed OT and the P-map as analytical tools for data analysis. The P-map is a mental representation of the degree of distinctiveness of different contrasts in various positions. In the OT system, the P-map serves as a mechanism that relates rankings between correspondence constraints to perceived differences of similarity degree. The crucial role of the P-map in the sound system is that it performs the function of guiding the speaker in search of the relatively minimal input deviation that solves a phonotactic problem. In what follows, I will present experimental results from Korean speakers' production of English consonant clusters and discuss two main patterns of vowel insertion identified in the data.

3. Production experiment

3.1 Experimental design

3.1.1 Participants

Fourteen EFL students in Korea participated in this experiment.[5] At the time of the experiment, all the participants were taking the class provided by English language and literature department in the university. Participants' ages ranged from 21 to 27. Participants who stayed in English-speaking countries more than three months were excluded from the study to keep the amount of English exposure similar. In the questionnaire about their English proficiency, the participants marked themselves either as "low middle" or "middle".

3.1.2 Materials

Nonsense words containing legal consonant clusters in English were used as stimuli to systematically capture a range of phonetic and phonological features in the stimuli. As a singleton onset and coda consonant, /t/ and /n/ were used, and as vowels, /e/ and /o/ were used. Mid vowels were chosen

[5] The total number of participants in the experiment was fifty. However, due to time constraints, fourteen participants' data, which were chosen based on their strong accents, were analyzed for the current study. Participants with strong accents were chosen based on the assumption that their production data may contain more errors.

instead of high vowels to avoid high vowels triggering a similar type of vowel insertion in consonant clusters by vowel harmony (e.g., Kim (2000) reported that Korean speakers insert /i/, /ʊ/, and /ɨ/ to resolve English consonant clusters). Refer to Appendix 1 for the wordlist used in the experiment. The types of clusters used in the experiment are presented in (7-8).

(7) Types of onset clusters
  a. voiced stop + liquid (#br, #bl, #dr, #gr, #gl)
  b. voiceless stop + liquid (#pr, #pl, #tr, #kr, #kl)
  c. strident + sonorant (#sn, #sm, #sl)
  d. strident + voiceless stop (#sk, #st, #sp)

(8) Types of coda clusters
  a. voiced stop + voiced obstruent (bd#, bz#, gd#, gz#)
  b. voiceless stop + voiceless obstruent (kt#, ks#, pt#, ps#)
  c. strident + stop (sp#, sk#, st#, zd#)
  d. sonorant + stop (lp#, lk#, lt#, ld#, lb#, mp#, nd#, nt#)
  e. sonorant + strident (ls#, ns#, nz#)
  f. liquid + nasal (lm#)

3.1.3 Procedure

In the practice session, participants were given a chance to read over the wordlist before being recorded so that they could familiarize themselves with the nonsense words. Each participant was instructed to read the word as it was spelled and was not allowed to ask questions on the pronunciation of a specific word. Participants were encouraged to read the nonsense words as if they were real English words.

Each word was presented to the participants on index cards one at a time, but pseudo-randomized order was kept to make sure there is no order effect. A filler, whose structure is VCCV, was inserted after every three words. The participants read each word embedded in a carrier phrase "I am saying___" The carrier phrase was used to provide consistent phonological environment before each target word. The experiment was recorded on a Marantz PMD 201 cassette recorder.

3.1.4 Measures

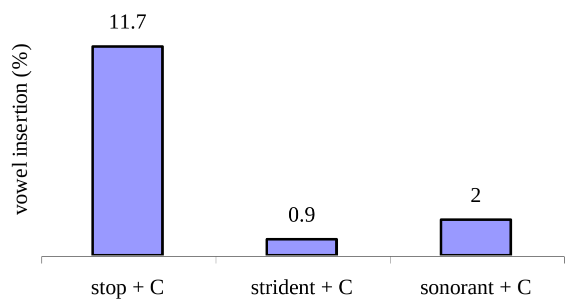
The recordings of Korean speakers' production of English consonant clusters were digitized through USB pre, and then waveforms and spectrograms were generated by Praat (version 4.2.04). First, the researcher made judgments on the tokens based on the spectrograms and waveforms. Following this, three native English speakers listened to the data and made their judgments on each token. The native English speakers were all graduate students in the linguistics department at Michigan State University.

An epenthetic vowel was identified with three criteria. First, vocalic complex patterns in the waveform were identified. Then, it was checked whether the vocalic complex patterns in the waveform were accompanied by clear F2 in the spectrogram. Lastly, the portion of a presumed epenthetic vowel was examined in an expanded view. If there were periodic wave shapes, the token was finalized as having an epenthetic vowel. Refer to Appendix 2 for the spectrogram and waveform containing an epenthetic vowel.

3.1.5 Results

Two major patterns of vowel insertion were identified in the experiment.[6] First, Korean speakers inserted a vowel more often in a cluster of stop + C than in a cluster of strident + C or sonorant + C. Figure 1 shows the vowel insertion pattern in the three cluster types.

Figure 1. The patterns of vowel insertion # 1



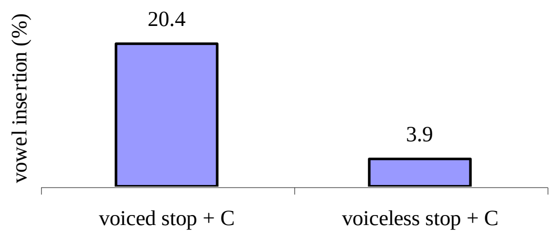
Among 1064 tokens of stop + C clusters, there were 125 tokens of vowel insertion (11.7%), whereas there are only 5 tokens of vowel insertion among 532 tokens of strident + C clusters (0.9 %). And among 784 tokens of sonorant + C clusters, there were 16 tokens of vowel insertion (2 %). In sum, we can clearly see that vowel insertion occurs most frequently when the first consonant of a cluster is a stop.

Interestingly, a closer look at the stop + C clusters reveals the asymmetry between a voiced stop and a voiceless stop: Korean speakers inserted a vowel more frequently in a cluster of voiced stop + C than in a cluster of voiceless stop + C. The following figure illustrates this pattern.

---

6 In the experiment, an epenthetic vowel was identified when it occurred between C_1 and C_2. (e.g., C_1C_2 #--> C_1VC_2 #). However, it is true that an epenthetic vowel can also occur after C_2 in coda position. In the current study, the vowel insertion after C_2 was not counted since the focus of this study is mainly on the epenthetic vowel which breaks down C_1 and C_2. The future research needs to improve upon this limitation.

Figure 2. The patterns of vowel insertion # 2



Among 504 tokens of voiced stop + C clusters, there were 103 tokens of vowel insertion (20.4 %). In contrast, among 560 tokens of voiceless stop + C clusters, there were only 22 tokens of vowel insertion (3.9 %). In particular, Korean speakers showed much difficulty with voiced stop + C coda clusters.[7]

Actually, these results are not new to L2 acquisition of consonant clusters. For instance, Kim (2000: 138) found that in Korean speakers' production of English consonant clusters, vowel epenthesis is more likely to occur in a cluster of voiced obstruent + C than in a cluster of voiceless obstruent + C. The same pattern was observed in Mandarin Chinese and Cantonese speakers' production of English consonant clusters (Chu 2005). Davidson also found that, given various consonant clusters that are illegal in English, English speakers are more likely to epenthesize a vowel in a cluster of voiced consonants than in a cluster of voiceless consonants (Davidson 2001: 44). In addition, Chen reported that Chinese learners of English in Taiwan made more mistakes in *p, t, k* + *r* clusters than in *s* + *p, t, k* clusters (Chen 2003: 9).

Then, the questions are a) why vowel insertion is more likely to occur after a stop than after a strident or a sonorant and b) among stop consonants, why vowel insertion is favored to occur after a voiced stop. In the following section, I provide possible answers to these questions based on the P-map proposal. I will also present P-map-projected-faithfulness constraints, their relevant rankings, and OT tableaux for each pattern of vowel insertion.

In particular, I present two proposals to account for the attested patterns of vowel insertion. First, following Côté (2000), I propose that the likelihood of vowel insertion in consonant clusters correlates with the perceptual salience of a segment in contrast: the weaker the perceptual cues of a segment, the more likely it is to be modified. In a similar vein, I also propose that L2 learners are guided by the P-map when faced with L2 phonotactic constraints that are not part of their native languages. That is,

[7] One of the frequent errors in coda clusters of voiced stop + C was devoicing of the first consonant. However, detailed analyses of this error type are not included in the present study. In addition, some of the Korean speakers in the experiment were not aware that the vowel in English becomes longer when preceded by a final voiced stop.

the P-map knowledge leads L2 learners to choose the output form with relatively minimal modification of the input in terms of perceptual similarity.

4. Proposal

4.1 The P-map and patterns of vowel insertion

4.1.1 Relative acoustic saliencies of a stop, a strident, and a sonorant

We have seen in the previous section that there occur asymmetries in the number of vowel insertions depending on the type of C_1 consonant in a cluster: a stop (11.7 %), a strident (0.9 %) or a sonorant (2 %). Regarding this asymmetry, I propose that the consonant whose acoustic cues are relatively weak (like stops) attracts vowel insertion, since it favors being closer to a vocalic segment to maintain its perceptibility in a string (Côté 2000: 131-132). By the same reasoning, the consonant whose acoustic cues are relatively strong can perceptually stand out well without a vocalic segment nearby.

To correlate the patterns of vowel insertion and the perceptual salience of a segment in a string, we need to look at the acoustic properties of consonants involved. First, a strident has relatively strong internal cues owing to its frication noise and hissing properties.[8] In addition, the inherent duration of a strident is especially long compared with other consonants: /s/ = 129 msec, /p/ = 89, /t/ = 77, /k/ =69, /b/ = 90, /d/ =83 /g/ = 67, /l/ = 66 (Umeda 1977: 851). The inherent duration is affected by many factors such as the following and preceding segment and stress, but the long duration of /s/ is nonetheless prominent among the consonants.

Like a strident, sonorant consonants such as a liquid and a nasal have relatively strong internal cues indicated by their formant structures. And therefore, they may not require additional vocalic segment to stand out in a string.

In contrast, the internal cues of a stop are weaker than other obstruents due to the complete blocking of airflow during its closure duration.[9] Therefore, a stop relies on the release burst to become audible. However, not all stops take advantage of this release burst: a release burst occurs in prevocalic position, but its occurrence is variable in post-vocalic position. Accordingly, non-prevocalic stops do not reliably benefit from a release burst, and the absence or weakness of the release burst may reduce the salience of stops (Côté 2000: 137). Therefore, a stop in stop + C cluster may require a vocalic segment to stand out in a string, especially when the second member of the cluster is an obstruent.

---

8 Internal cues are produced during the closure part of the consonant (Côté 2000: 133).

9 Acoustically, the closure duration of a stop is indicated by a weak transient (in the case of voiced stops) or a complete silence (in the case of voiceless stops).

Given the relative acoustic saliency differences in the three types of consonants (a stop, a strident, and a sonorant), we can see that there is correlation between the relative acoustic salience of a segment and the number of vowel insertions. As was discussed, a stop has relatively weak perceptual cues, and as a result, it may need a vocalic segment nearby to perceptually stand out in a string. Our data showed that vowel insertion occurs most frequently after a stop.

Putting it differently, we can interpret the relatively frequent occurrence of an epenthetic vowel after a stop as a result of articulatory effort to increase the perceptibility of a stop in a string. Given the relatively weak acoustic cues of a stop, it is reasonable to assume that an epenthetic vowel occurs as a by-product of L2 learners' effort to maintain the perceptibility of a stop in their production.

Recently, Yoo (2004) made a similar line of claims regarding the patterns of vowel insertion in the production of English consonant clusters by Korean children. The data in the study were collected from 18 children three times over the period of 3 years. Yoo noted that insertion rate changed over the period of three years (2.2% → 10.2% → 1.5%) and she interpreted insertion errors as "a typical simplification type of the lower level foreign language learner who wants to pronounce all the consonants they hear or read, by making unmarked CV syllables" (Yoo 2004: 491).[10] It seems that Yoo's interpretation of vowel insertion error can be restated as L2 learners' effort to make their production perceptible.

4.1.2 Relatively minimal deviation from input

In the previous section, I have proposed that vowel epenthesis helps to improve the perceptibility of a stop in consonant clusters. In this section, I account for the observation that an epenthetic vowel occurs frequently after a stop (in particular, after a voiced stop), based on the key concept of the P-map such that a sound change whose modification is less perceptible is more likely to occur than a sound change whose modification is more perceptible.

Applying the concept of the P-map to the first pattern of vowel insertion (the asymmetry of a stop, a strident, and a sonorant), we can say that vowel insertion occurs more often after a stop, because an epenthetic vowel after a stop brings about less perceptual deviation from the unepenthesized input than does an epenthetic vowel after a strident or a sonorant. For instance, the production results indicate that, if the P-map proposal is correct, the epenthesized output $VC_{[stop]}$**ɨ**C# is a lesser deviation from the input

10 Yoo (2004) reported that vowel insertion occurred in the following onset clusters in her study: /sk, gr, dr, pr/ in 2001, /bl, gl, gr, sk/ in 2002, and /bl, gl, gr, br/ in 2003. Since she did not provide the number of vowel insertions in each cluster, it is not clear whether there occurred more vowel insertion after a stop than after a strident. However, at least, it seems that vowel insertion involves more stop consonants than a strident.

$VC_{[stop]}C\#$ than $VC_{[strident]}ɨC\#$ is of the input $VC_{[strident]}C\#$ or $VC_{[sonorant]}ɨC\#$ is of the input $VC_{[sonorant]}C\#$. What follows is the justification for the degree of deviation in the two pairs, $VC_{[stop]}ɨC\#$ -$VC_{[stop]}C\#$ vs. $VC_{[strident]}ɨC\#$ - $VC_{[strident]}C\#$.

First, the hypothesized P-map for vowel epenthesis proposed by Steriade (2001b) shows that vowel epenthesis in a cluster of strident + obstruent may cause serious dissimilarity between unepenthesized input and epenthesized output.

(9) P-map on vowel epenthesis (adapted from Steriade 2001b: 18)
Bigger letter = less similar to unepenthesized input

Zero/vowel	T_Ri	S_T
Ø / ə	Ø / ə	Ø / ə
Ø / u	Ø / u	Ø / U
Ø / i	Ø / i	Ø / i

In Table (9), "T" stands for an obstruent, "R" stands for a sonorant, and "S" stands for a strident. Notation Ø /u in each cell indicates the perceptual difference between the unepenthesized input (Ø) and vowel epenthesized output (ə, u, i).

Differences in the size of the notations (e.g., Ø /u vs. Ø /U) indicate the perceptual difference between input and output: the bigger the size of the letters, the less similar the epenthesized output is to unepenthesized input. For instance, compare the size of Ø/u, Ø/ə and Ø/i under the column of T_Ri and S_T. We can see that the letters under the column of S_T are bigger than those under the column of T_Ri. This is interpreted such that epenthetic vowels /u, i, ə/ in the environment of S_T make the epenthesized output less similar to the unepenthesized input than those in the environment of T_Ri. In short, Table (9) shows that vowel insertion in the environment of S_T results in a greater perceptual difference from the unepenthesized input, and therefore is less preferred cross-linguistically.

Fleischhacker's research testing perceptual similarity between consonant cluster input and vowel epenthesized output demonstrates that the hypothesized P-map (Table 9) is actually correct in predicting where vowel insertion is likely to occur. In her study, Fleischhacker (2001) tested the similarity between two sets of stimuli, VSC–SC vs. SVC-SC (V = epenthetic vowel, S = strident, C = consonant). The results showed that VSC-SC stimuli sound more similar to each other than SVC-SC stimuli, while the reverse is true in the case of obstruent + sonorant (Fleischhacker 2001: 29). This research demonstrates that vowel insertion between voiceless strident + stop (as in SVC-SC stimuli) is not desirable in terms of preserving perceptual similarity to the unepenthesized input.

Interestingly, a similar pattern was revealed in the author's experiment on perceptual similarity of unepenthesized input and epenthesized output, which involved 38 Korean speakers. An epenthetic vowel after a strident brings about much more perceptual difference from the unepenthesized input than an epenthetic vowel after a voiced stop does. More detailed discussion of this experiment will be presented in Section 5.

Another finding in this study is that no sound modifications occurred in the onset cluster of strident + stop (e.g., *skot*, *stet*). It seems that onset [s] + stop clusters are especially easy for L2 learners to acquire. Actually, we can find supporting evidence for this observation in Morelli's work, in which she proposed that [s]+ stop clusters are unmarked within the domain of obstruent clusters (Morelli 2003).

As a response to the long-standing question of why [s] + stop onset clusters frequently occur across languages, even though [s] + stop apparently violates the sonority sequencing principle, Morelli (2003) provides evidence that [s] + stop clusters are unmarked (among obstruent clusters) in terms of both place and manner dimensions. That is, [s] + stop clusters do not violate manner related constraints, disallowing onsets with tautosyllabic continuant segments, tautosyllabic non-continuant segments, and stop + obstruent sequences. On the dimension of place features, [s] + stop clusters are unmarked in that the least marked place of articulation (i.e., coronal) occurs in the articulation of the first obstruent which is in a position of weak perceptibility.[11]

The unmarkedness of [s] + stop clusters are also found in child language acquisition. It has been shown that, while most children start with obstruent-sonorant clusters, some children first have /s/-obstruent clusters. For instance, in a longitudinal study on the acquisition of Portuguese syllables, Freitas found that children first acquire /s/-obstruent clusters, despite the fact that Portuguese has both obstruent-sonorant and /s/-obstruent clusters (Freitas 1997).

The asymmetry of vowel insertion between stop + C, strident + C and sonorant + C clusters is captured through the following OT constraints.

(10) Faithfulness constraints for the asymmetry among a stop, a strident, and a sonorant

a. DEP-V/ strident C_1 _C_2: A vowel present in the output context of strident C_1_C_2 has a correspondent vowel in the input context of strident C_1_C_2.

b. DEP-V/ stop C_1_C_2: A vowel present in the output context of stop C_1_C_2 has a correspondent vowel in the input context of stop C_1_C_2.

c. DEP-V/ sonorant C_1_ C_2: A vowel present in the output context of sonorant C_1_C_2 has a correspondent vowel in the input context of sonorant C_1_C_2.

d. Ranking Relation: DEP-V/ strident C_1 _C_2, DEP-V/ sonorant C_1_C_2 >> DEP-V/ stop C_1_C_2

[11] The consonant in preconsonantal position is less likely to be perceived since it can be masked by the presence of a following consonant.

Note that the faithfulness constraints in (10) are projected by the P-map. We have discussed in Section 2.2 that based on the perceptual similarity between the input and output form, the P-map projects faithfulness constraints and determines its ranking in the OT system. For instance, if a contrast x-y/_A is more perceptible (or distinctive) than a contrast x-z/_A, then the P-map's effect on the grammar will rank higher the correspondence constraint x-y/_A than correspondence x-z/_A.

Regarding the justifications for the constraint ranking in (10, d), we have discussed two properties of consonants at issue: a) relative acoustic saliencies of a stop, a strident, and a sonorant in Section 4.1.1, and b) relatively minimal modification of the input in Section 4.1.2. The ranking relation also reflects the experimental finding that vowel insertion is more likely to occur in a cluster of stop + C than in a cluster of strident + C or in a cluster of sonorant + C. Tableaux (11) and (12) illustrate the interaction among these constraints.

(11) /tesk/ → [tesk]

/tesk/	MAX-IO	DEP-V/ strident $C_1\_\ C_2$	DEP-V/ sonorant $C_1\_\ C_2$	*COMPLEX CODA	DEP-V/stop $C_1\_\ C_2$
☞a. tesk				*	
b. tesɨk		*!			
c. tes	*!				
d. tek	*!				

(12) /tebd/ → [tebɨd]

/tebd/	MAX-IO	DEP-V/ strident $C_1\_\ C_2$	DEP-V/ sonorant $C_1\_\ C_2$	*COMPLEX CODA	DEP-V/stop $C_1\_\ C_2$
a. tebd				*!	
☞b. tebɨd					*
c. teb	*!				
d. ted	*!				

Tableau (11) illustrates the case where DEP-V/ strident $C_1\_\ C_2$ is ranked higher to prevent vowel insertion in a cluster of strident + C, and consequently, no vowel insertion occurs. On the other hand, tableau (12) shows that DEP-V/stop $C_1\_C_2$, the constraint preventing vowel insertion after a stop, is ranked lower than *COMPLEX so that vowel insertion after a stop is allowed as an optimal output. In particular, the ranking relation, DEP-V/ strident $C_1\_\ C_2$, DEP-V/ sonorant C1_C2 >> *COMPLEX >> DEP-V/stop $C_1\_C_2$, illustrates the case where the same speaker inserts a vowel variably depending on the cluster type. For instance, participant 10, who produced s + C clusters correctly, inserted a vowel after a stop.

4.1.3 Asymmetry between a voiced stop and a voiceless stop

In the previous section, we have seen that vowel insertion is most likely to occur after a stop. Moreover, a closer look at the number of vowel insertions in a cluster of stop + C reveals that more vowel insertion occurred after a voiced stop (20. 4 %) than after a voiceless stop (3.9 %). In this section, I deal with the question of why vowel insertion is more likely to occur after a voiced stop than after a voiceless stop.

There has been consistent evidence that vowel epenthesis is more tolerable and, in some sense, beneficial after a voiced stop than after a voiceless stop. According to Côté (2000: 167), the tendency to insert a vowel after a voiced stop results from the fact that voiced segments share the low frequency energy associated with voicing of vowels, and therefore vowel epenthesis becomes less noticeable in the context of voiced segments. Similarly, Fleischhacker (2001) argues that a voiced environment is more similar to a vocalic element than a voiceless environment, and consequently epenthesis is more tolerable in the former than in the latter. In sum, the findings of Côté (2000) and Fleischhacker (2001) suggest that vowel insertion is more likely to occur after a voiced stop, because an epenthetic vowel after a voiced stop minimizes the perceptual dissimilarity to the unepenthesized input by virtue of the fact that a voiced stop and an epenthetic vowel share vocalic properties. I tested this proposal by conducting an auditory similarity judgment experiment between unepenthesized input and epenthesized output. The detailed discussion of this experiment is presented in Section 5.

The following P-map projected faithfulness constraints are proposed to capture the asymmetry of vowel insertion between a voiced stop and a voiceless stop.

(13) Faithfulness constraints for the asymmetry between a voiced stop vs. a voiceless stop.
- a. DEP-V/ voiceless stop $C_1\_C_2$: A vowel present in the output context of voiceless stop $C_1\_C_2$ has a correspondent vowel in the input context of voiceless stop $C_1\_C_2$.
- b. DEP-V/ voiced stop $C_1\_ C_2$: A vowel present in the output context of voiced stop $C_1\_C_2$ has a correspondent vowel in the input context of voiced stop $C_1\_C_2$.
- c. Ranking relation: DEP-V/ voiceless stop $C_1\_C_2$ >> DEP-V/ voiced stop $C_1\_ C_2$

Again, the ranking relation DEP-V/ voiceless stop $C_1\_C_2$ >> DEP-V/ voiced stop $C_1\_ C_2$ finds its justification in Côté (2000) and Fleischhacker (2001) which showed that vowel insertion after a voiced stop is perceptually less obtrusive. Tableau (14) illustrates the interaction of the constraints in (13).

(14) /tegd/ → [tegɨd]

/tegd/	MAX-IO	DEP-V/ strident C1_ C2	*COMPLEX CODA	DEP-V/ voiceless $C_1\_C_2$	DEP-V/ voiced $C_1\_C_2$
a. tegd			*!		
☞ b. tegɨd					*
c. tekd			*!		
d. tekɨd				*!	
e. tekt			*!		
f. tekɨt				*!	
g. teg	*!				
h. ted	*!				

Tableau (14) demonstrates the case where a voiced stop + C cluster is realized with an epenthetic vowel. The ranking relation between DEP-V/ voiceless stop $C_1\_ C_2$ >> DEP-V/ voiced stop $C_1\_ C_2$ reflects the observation that vowel insertion occurs more frequently in voiced stop + C clusters than in voiceless stop + C clusters. For instance, both [tekɨd] and [tekɨt] lose their candidacies since they violate highly ranked constraint DEP-V/ voiceless stop $C_1\_C_2$.

To summarize, in this section, vowel insertion patterns in onset and coda clusters were discussed. Two major patterns of vowel insertion were observed. First, vowel insertion is more likely to occur in a cluster of stop + C than in a cluster of strident + C or sonorant + C. Second, vowel insertion is more likely to occur in a cluster of voiced stop + C than in a cluster of voiceless stop + C.

Regarding the asymmetry among a stop, a strident, and a sonorant, I found the perceptual justification of the attested patterns in the relative acoustic saliency-differences among segments. A stop has relatively weak acoustic cues compared to a strident or a sonorant. L2 learners are also sensitive to these acoustic differences, and accordingly an epenthetic vowel after a stop occurs as a by-product of an articulatory effort to increase the perceptibility of a stop in a string. In addition, Fleischhacker (2001) showed that vowel insertion between voiceless strident + stop is not desirable in terms of preserving perceptual similarity to the unepenthesized input.

When it comes to the asymmetry between a voiced stop and a voiceless stop, it is suggested that an epenthetic vowel shares a vocalic environment with a voiced stop, and therefore an epenthetic vowel after a voiced stop is less obtrusive than that in voiceless environment. The constraint rankings regarding vowel insertion can be summarized as follows.

(15) DEP-V constraints (Summary)

a. DEP-V/ strident $C_1 \_C_2$, DEP-V/ sonorant $C_1\_ C_2$>> DEP-V/ stop $C_1\_C_2$

b. DEP-V/ voiceless stop $C_1\_C_2$ >> DEP-V/ voiced stop $C_1\_ C_2$.

5. Auditory similarity experiment

5.1 Experimental design

In this section, I provide more direct evidence for the proposals in the study that a) the likelihood of vowel insertion in consonant clusters correlates with the perceptual salience of a segment in contrast: the weaker the perceptual cues of a segment, the more likely it is to be modified, and b) the P-map knowledge leads L2 learners to choose the output form with relatively minimal modification of the input in terms of perceptual similarity.

Based on the production results discussed in Section 3 and Section 4, it is hypothesized that a) Korean speakers judge the unepenthesized consonant clusters and epenthesized output as perceptually more similar to each other a) when C_1 is a voiced stop rather than when C_1 is a strident (e.g., [tebd]-[tebʊd] is more similar than [test]-[tesʊt]) and b) when C_1 is a voiced stop rather than when C_1 is a voiceless stop (e.g., the pair [tebd]-[tebʊd] is more similar than [tept]-[tepʊt]).

This is an interesting proposal if proven, considering that Korean speakers may not reliably judge the presence of an epenthetic vowel in a cluster. Korean speakers may hear an illusionary vowel in consonant clusters while complying with a Korean language-specific phonotactic constraint that does not allow complex margins (as is the well-known case with Japanese speakers in Dupoux et al. 1998). Besides, Korean speakers may not notice that the difference between the pair [bret] and [bʊret] is the existence of an epenthetic vowel that breaks up the original cluster.

Therefore, whether Korean speakers respond differently to consonant clusters with an epenthetic vowel depending on cluster types (i.e., whether they think an epenthetic vowel in a certain cluster more perceptible or less perceptible) is an interesting question, considering Korean speakers' impoverished perception of an epenthetic vowel in a cluster.

An auditory similarity judgment experiment was designed to collect Korean EFL learners' judgments on perceptual similarity between unepenthesized input and epenthesized output. In the experiment, the degree of similarity to the input for each set of epenthesized outputs is tested.

5.1.1 Materials

The patterns of vowel insertion identified in the production experiment involve two distinctive features of C_1 in the clusters: a) a voicing feature of C_1 (voiced vs. voiceless), and b) a manner feature of C_1 (stop vs. strident). To incorporate these features in the auditory similarity judgment test, three groups of stimuli were generated: a) a voiced stop + C vs. a strident + C, b) a voiceless stop + C vs. a strident + C, and c) a voiced stop + C vs. a voiceless stop + C.

The stimuli for the experiment were nonsense words containing legal English consonant clusters and their associated pairs of epenthesized forms. Nonsense words were used to systematically capture a range of phonetic and phonological features in the stimuli. Also, using nonsense words makes it possible to avoid similarity judgments being affected by the participants' familiarity with the real words in the stimuli (see Flege et al. 1996 for the discussion of the influence of lexical familiarity in L2 perception).

English [ʊ] is chosen as an epenthetic vowel, which is assumed to be the closest approximation of the Korean vowel [ɨ] (Kabak 2003: 72). As a single onset and coda, [n], a legal onset and coda in Korean, was chosen. The stimuli were recorded by a male native speaker of English who was naive to the purpose of the experiment. This speaker read the word list, which is composed of nonsense words such as *buREN* and *BREN*. The speaker was instructed to put a stress on the capital letter of the word in the wordlist. In this way, the first vowel in a word like *buREN* was realized as high, back, lax vowel, [ʊ].

Each cluster was paired with its associated epenthesized form (e.g., bren-buren). Every utterance was followed by a one-second pause. One-second pause was judged to be appropriate for conducting a test by the participants in the pilot study. And then two pairs of sounds were arranged in a sequence for the purpose of similarity comparison among groups (e.g., voiced group vs. voiceless group = bren-buren vs. pren-puren). The order of the two groups in each template was mixed so that the participants' judgment was not biased by the sequence of the stimuli (e.g., template A: dren-duren vs. sten-suten, template B: spin-supin vs. drin vs. durin).

In summary, each item in the stimuli contains four sounds in a sequence (A-B vs. C-D) and participants were asked whether the pair A-B is more similar to each other or the pair C-D is more similar to each other. The stimuli by group are given in Table 16. The epenthesized form in each stimulus was not presented in Table 16 for simplicity's sake. Thus, the actual template for each stimulus looks like, for instance, /dren-dʊren-sten-sʊten/.

(16) Stimuli

voiced stop-strident	voiceless stop-strident	voiced-voiceless
dren-sten	tren-sten	dren-pren
spun-brun	spun-prun	plin-glin
brin-stin	prin-stin	bren-cren
bren-sken	pren-sken	clun-drun
spin-drin	spin-trin	nikt-nigd
smen-glen	smen-clen	nebz-neps

snun-brun	snun-plun	nubd-nupt
slin-grin	slin-crin	nigz-niks
nigd-nisp	nikt-nisp	nugd-nukt
nisk-nibz	nisk-nips	negd-nekt
nesk-nebz	nesk-neps	negz-neks
nigd-nist	nikt-nist	
nest-nebd	nest-nept	
nusp-nugd	nusp-nukt	
nubz-nust	nups-nust	
negd-nesp	nekt-nesp	

5.1.2 Procedure

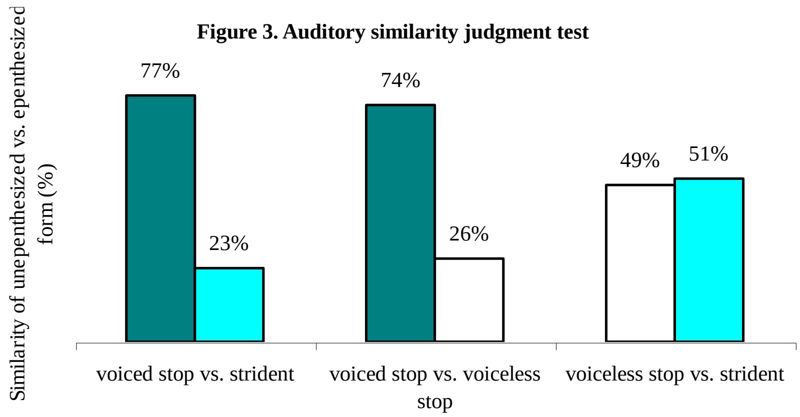
Thirty-eight native speakers of Korean participated in the study. The researcher met with each participant in a quiet room. The participant used a headphone (Samsung smh-m150), and the stimuli were played over a desktop computer (TG dreamsys/R3LC, Pentium 4). Participants were instructed to make their judgments on the auditory similarity of two pairs of words, i.e., the similarity between an unepenthesized form and its associated epenthesized form. They were asked to circle the pair that sounded more similar to each other.

The recorded lists were accompanied by an answer sheet presenting only the unepenthesized forms. The epenthesized forms were not presented in the answer sheet so that participants were not aware of the fact that each item is modified in a systematic way by inserting an epenthetic vowel. The type of modification is disguised so as to avoid a situation in which the participants pay too much attention to an epenthetic vowel itself rather than hearing the whole words when making their judgments on the similarity of the pairs. A pretest containing six stimuli was given to each participant to check whether the participant understood the test procedure. In the experiment, no participants were reported as having difficulty in conducting the test.

5.1.3 Results

The stimuli were presented in two randomly ordered lists and results from both ordered lists were pooled in Figure 3. Figure 3 contains three groups of bar graphs: a) a voiced stop + C vs. a strident + C, b) a voiced stop + C vs. a voiceless stop + C, and c) a voiceless stop + C vs. a strident + C. Two bar graphs in each group make up 100 % of the responses for each category.

Figure 3. Auditory similarity judgment test



* X-axis is abbreviated for simplicity, and each color of the bars represents the following.

	the case where voiced stop + C is paired with voiced stop + epenthetic vowel + C
	the case where voiceless stop + C is paired with voiceless stop + epenthetic vowel + C
	the case where strident + C is paired with strident + epenthetic vowel + C

The first group of bar graphs is the responses on the stimuli with voiced stop + C vs. strident + C (e.g., bren_buren vs. sten_suten). The dark gray bar is the percentage of responses on voiced stop + C and its associated epenthesized form (77%), and the light gray bar is the percentage of the responses on the strident + C and its associated epenthesized form (23%). Stated another way, 77% of the participants said that voiced stop + C and its associated epenthesized form sound more similar, while 23% said that strident + C and its epenthesized form are more similar.

The graph in Figure 3 shows that Korean speakers dominantly judged voiced stop + C and its epenthesized form as more similar to each other than strident + C and its epenthesized form. Also, the second group of bar graphs shows that when voiced stop + C is paired with voiceless stop + C, Korean speakers judge voiced stop + C and its epenthesized form as more similar to each other than voiceless stop + C and its epenthesized form (74% vs. 26%). However, when voiceless stop + C is paired with strident + C, the similarity judgment is almost evenly distributed as is shown in the third group of bar graphs (49% vs. 51%).

In sum, the experimental results clearly demonstrate that an epenthetic vowel after a voiced stop is perceptually less obtrusive to Korean speakers than an epenthetic vowel after a strident or a voiceless stop. Therefore, the proposal that vowel insertion is more likely to occur after a voiced stop because the epenthetic vowel after a voiced stop incurs less deviation from unepenthesized input were successfully supported by the auditory similarity judgment test. The fact that the auditory similarity judgment is evenly distributed when voiceless stop + C is paired with strident + C requires further investigation, because in the production data, there were more vowel insertions after a voiceless stop than after a strident.

6. Conclusion

This study investigated the patterns of vowel insertion in Korean speakers' acquisition of English consonant clusters by implementing both a production and a perception experiment. In the production experiment, two major patterns of vowel insertion were identified: a) vowel insertion is more likely to occur after a stop than after a strident or a sonorant, and b) vowel insertion is more likely to occur after a voiced stop than after a voiceless stop.

Following the P-map proposal, it was proposed in this study that L2 learners choose the output form with relatively minimal modification of the input in terms of perception, while resolving phonotactic constraints in L2. The patterns of vowel insertion in Korean speakers' production of English consonant clusters clearly demonstrated that the site of vowel insertion is chosen under the guidance of the P-map.

To provide more direct evidence for the proposal above, the degree of similarity to unepenthesized input for each set of epenthesized outputs was tested. The auditory similarity judgment test successfully showed that vowel insertion after a voiced stop is less obtrusive than vowel insertion after a strident or a voiceless stop. Consequently, we can safely say that Korean speakers' production of English consonant clusters are guided by the P-map principle which states that a sound change whose modification is less perceptible is more favored than a sound change whose modification is more perceptible.

To conclude, I have shown in this study that Korean speakers' production of English consonant clusters is by no means random, instead it follows a certain principle: Deal with the phonotactically illegal structures, but do it in a minimally obtrusive way.

References

ABRAHAMSSON, NICLAS. 2003. Development and recoverability of L2 codas: A longitudinal study of Chinese-Swedish interphonology. *Studies in Second Language Acquisition* 25, 313-49.

AHN, SANG-CHEOL. 1985. *The interplay of phonology and morphology in Korean*. Ph.D dissertation. University of Illinois.

ANDERSON, JANET I. 1987. The markedness differential hypothesis and syllable structure difficulty. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 279-91. Cambridge, MA: Newbury House.

BENSON, B. 1988. Universal preference for the open syllables as an independent process in interlanguage phonology. *Language Learning*

38, 221-35.

BROSELOW, ELLEN. 1987. Non-obvious transfer: On predicting epenthesis errors. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 292-304. Cambridge, MA: Newbury House.

BROSELOW, ELLEN, and DANIEL FINER. 1991. Parameter setting in second language phonology and syntax. *Second Language Research* 7, 35-59.

CHEN, SZU-WEI. 2003. Acquisition of English onset clusters by Chinese learners in Taiwan. Paper presented at Postgraduate Conference 2003, University of Edinburgh.

CHU, G.Y.-F. 2005. Clusters in L2 phonology: Evidence from Mandarin and Cantonese. Paper presented at 1st ASA Workshop on Second Language Speech Learning, Vancouver, BC.

CÔTÉ, MARIE-HÉLÈNE. 2000. *Consonant cluster phonotactics: A perceptual approach*. Ph.D dissertation. MIT.

DAVIDSON, LISA. 2001. Hidden rankings in the final state of the English grammar. *RuLing Papers II*, ed. by G. Horwood and S.-K. Kim. New Brunswick: Rutgers University.

DUPOUX, E, K KAKEHI, Y HIROSE, C PALLIER, and J MEHLER. 1998. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25, 1568-78.

ECKMAN, FRED R. 1981a. On the naturalness of interlanguage phonological rules. *Language Learning* 31, 195-216.

__________________. 1981b. On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition* 4, 18-30.

ECKMAN, FRED R., and GREGORY K. IVERSON. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9, 234-52.

FAN, YE. 2004. Modification of English complex codas by Mandarin Speakers: An Optimality Theory and P-map account. Ms. Michigan State University.

FLEGE, JAMES EMIL, NAOYUKI TAKAGI, and VIRGINIA MANN. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *The Journal of the Acoustical Society of America* 99, 1161-73.

FLEISCHHACKER, HEIDI. 2001. Cluster-dependent epenthesis asymmetries. UCLA working papers in linguistics 7: Papers in phonology 5.

FREITAS, M. J. 1997. *Aquisição da Estrutura Silábica do Português Europeu*. Ph.D dissertation. Universidade de Lisboa.

GREENBERG, JOSEPH, and J. J. JENKINS. 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20, 157-77.

GUI, CAN-KUN. 1985. *Mei Guo Ying Yu Ying Yong Yu Yin Xue [Applied phonetics of American English]*. Shanghai: Shanghai Foreign Language

Education Press.
HANCIN-BHATT, BARBARA, and RAKESH M. BHATT. 1997. Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition* 19, 331-78.
HANSEN, JETTE G. 2001. Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics* 22, 338-65.
KABAK, BARIS. 2003. *The perceptual processing of second language consonant clusters*. Ph.D dissertation. University of Delaware.
KAGER, RENE. 1999. *Optimality Theory*. Cambridge, United Kingdom: Cambridge University Press.
KIM, CHIN W., and HYOUNG-YOUB KIM. 1991. The Character of Korean glides. *Studies in the Linguistic Sciences* 21, 113-25.
KIM, H-Y. 1990. *Voicing and tensification in Korean: A multi-face approach*. Ph.D dissertation. University of Illinois.
KIM, HYOUK-KEUN. 2000. *The interlanguage phonology of Korean learners of English: A computational implementation of optimality theoretic constraints*. Ph.D dissertation. Georgetown University.
LADO, ROBERT. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
LEE, J-S. 1991. Metrical foot and vowel length in verbal morphology in Korean. *Harvard Studies in Korean linguistics* IV, 133-42.
LEE, YONGSUNG. 1993. *Topics in the Vowel Phonology of Korean*. Ph.D dissertation. Indiana University.
MAJOR, ROY C., and MICHAEL C. FAUDREE. 1996. Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition* 18, 69-90.
MCCARTHY, JOHN J. 1995. Extensions of faithfulness: Rotuman Revisited. Ms. University of Massachusetts, Amherst.
MCCARTHY, JOHN J., and ALAN S. PRINCE. 1995. Faithfulness and reduplicative identity. *Papers in Optimality Theory*, 249-384. Amherst Mass: Graduate Linguistic Student Association.
MORELLI, FRIDA. 2003. The relative harmony of /s+stop/ onsets: Obstruent clusters and the sonority sequencing principle. *The syllable in optimality theory*, ed. by Caroline Fery and Ruben van de Vijver, 356-71. New York: Cambridge U Press.
PRINCE, ALAN, and PAUL SMOLENSKY. 1993. Optimality theory: Constraint interaction in generative grammar. Ms. Rutgers University and University of Colorado at Boulder.
SATO, CHARLENE J. 1984. Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning* 34, 43-57.
________________. 1989. Task variation in interlanguage phonology. *Linguistic Perspectives on second language acquisition*, ed. by S. M. Gass and J. Schachter. New York: Cambridge University Press.

SOHN, HYANG-SOOK. 1987. *Underspecification in Korean Phonology*. Ph.D dissertation. University of Illinois.

STERIADE, DONCA. 2001a. Directional asymmetries in place assimilation: A perceptual account. *The role of speech perception in phonology*, ed. by E. Hume and K. Johnson. San Diego: Academic Press.

_______________. 2001b. The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms. UCLA.

TARONE, ELAINE E. 1987. Some influences on the syllable structure of interlanguage phonology. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 232-47. Cambridge, MA: Newbury House.

TROPF, HERBERT S. 1987. Sonority as a variability factor in second language phonology. *Sound patterns in second language acquisition*, ed. by Allan James and Jonathan Leather, 173-91. Dordrecht: Foris.

UMEDA, NORIKO. 1977. Consonant duration in American English. *Journal of the Acoustical Society of America* 61, 846-58.

VAN DEN BROECKE, M. P. R. 1976. *Hierarchies and rank orders in distinctive features*. Assen/Amsterdam: Van Gorcum.

VITZ, PAUL, and BRENDA SPIGEL WINKLER. 1972. Predicting the judged "Similarity of Sound" of English words". *Journal of Verbal Learning and Verbal Behavior* 12, 373-88.

WALDEN, BRIAN, and MONTGOMERY. 1975. Dimensions of consonant perception in normal and hearing impaired listeners. *Journal of Speech and Hearing Research* 18, 445-55.

WEINBERGER, STEVEN. 1987. The influence of linguistic context on syllable simplification. *Interlanguage phonology: The acquisition of a second language sound system*, ed. by Georgette Ioup and Steven Weinberger, 401-17. Cambridge, MA: Newbury House.

YOO, HYEBAE. 2004. A longitudinal study of consonant cluster acquisition. *Studies in Phonetics, Phonology and Morphology* 10, 481-503. The Phonology-Morphology Circle of Korea.

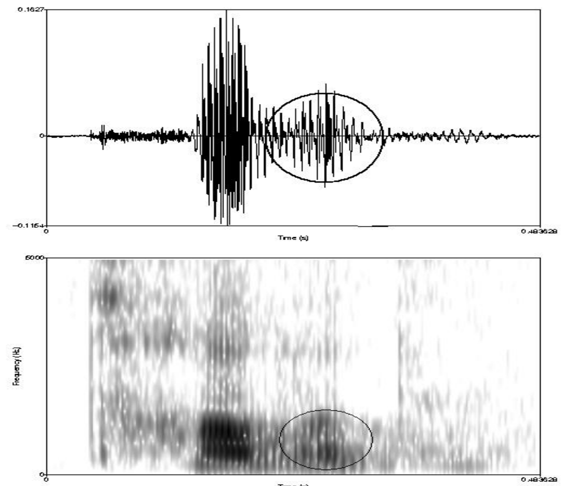
Appendix A. A Wordlist in the Production Experiment

Target Cluster	Attested Word	Target Stimuli (Nonsense words)	
[pr]	Prawn	Pret/Prot	Pren/Pron
[br]	Brown	Bret/Brot	Bren/Bron
[tr]	Train	Tret/Trot	Tren/Tron
[dr]	Drain	Dret/Drot	Dren/Dron
[gr]	Green	Gret/Grot	Gren/Gron
[kr]	Cream	Cret/Crot	Cren/Cron
[gl]	Glad	Glet/Glot	Glen/Glon
[kl]	Claim	Clet/Clot	Clen/Clon
[pl]	Play	Plet/Plot	Plen/Plon
[bl]	Blame	Blet/Blot	Blen/Blon
[sl]	Slim	Slet/Slot	Slen/Slon
[sn]	Snow	Snet/Snot	Snen/Snon
[sm]	Smile	Smet/Smot	Smen/Smon
[st]	Star	Stet/Stot	Sten/Ston
[sk]	Sky	Sket/Skot	Sken/Skon
[sp]	Sphere	Spet/Spot	Spen/Spon
[kt]	Act	Tekt/Tokt	Nekt/Nokt
[pt]	Adopt	Tept/Topt	Nept/Nopt
[bd]	Tabed	Tebd/Tobd	Nebd/Nobd
[gd]	Taged	Tegd/Togd	Negd/Nogd
[gz]	Tags	Tegz/Togz	Negz/Nogz
[ks]	Parks	Teks/Toks	Neks/Noks
[ps]	Lapse	Teps/Tops	Neps/Nops
[tz]	Spitz	Tetz/Totz	Netz/Notz
[bz]	Tabs	Tebz/Tobz	Nebz/Nobz
[sp]	Lisp	Tesp/Tosp	Nesp/Nosp
[sk]	Risk	Tesk/Tosk	Nesk/Nosk
[st]	List	Test/Tost	Nest/Nost
[zd]	Cruised	Tezd/Tozd	Nezd/Nozd
[nt]	Ant	Tent/Tont	Nent/Nont
[ns]	Rinse	Tens/Tons	Nens/Nons
[nd]	Tend	Tend/Tond	Nend/Nond
[ns]	Lens	Tenz/Tonz	Nenz/Nonz
[mp]	Lamp	Temp/Tomp	Nemp/Nomp
[lp]	Help	Telp/Tolp	Nelp/Nolp
[lb]	Bulb	Telb/Tolb	Nelb/Nolb
[lt]	Belt	Telt/Tolt	Nelt/Nolt
[ld]	Held	Teld/Told	Neld/Nold
[lm]	Film	Telm/Tolm	Nelm/Nolm
[ls]	Pulse	Tels/Tols	Nels/Nols

Appendix B. A Waveform and Spectrogram with an Epenthetic Vowel

Figure 4 presents a Korean speaker's production of a target word /tobd/. Epenthetic vowels are indicated by circles.

Figure 4. A waveform and spectrogram with an epenthetic vowel



Bo-Young Kwon
Department of Linguistics and Germanic, Slavic, Asian and African Languages
Michigan State University
A-614 Wells Hall East Lansing, MI 48824-1027
Phone: 1-517-353-0740
Fax: 1-517-432-2736
E-mail: kwonboyo@msu.edu

Received: July 15, 2005
Accepted: August 29, 2005