Explaining Phonetic Variation of Consonants in Vocalic Context

Eunjin Oh*

ABSTRACT

This paper aims to provide preliminary evidence that (at least part of) phonetic phenomena are not simply automatic or arbitrary, but are *explained* by the functional guidelines, ease of articulation and maintenance of contrasts. The first study shows that languages with more high vowels (e.g., French) allow larger consonantal deviation from its target than languages with less high vowels (e.g., English). This is interpreted as achieving the economy of articulation to a certain extent in order to avoid otherwise extreme articulatory movement to be made in CV syllables due to strict demand on maintaining vocalic contrasts. The second study shows that Russian plain bilabial consonant allows less amount of undershoot due to the neighboring vowels than does English bilabial consonant. This is probably due to the stricter demand on maintaining the consonantal contrasts, plain vs. palatalized, existing only in Russian.

Keywords: Phonetics, Phonology, Degree of Undershoot, Ease of Articulation, Maintenance of Contrasts

1. Introduction

The velar consonant in *key* [ki] is produced farther forward than the same consonant in *coo* [ku] in anticipation of the following front vowel. This kind of phonetic variation of a segment due to adjacent sounds has often been assumed to be an automatic realization of the phonetic implementation rule, e.g., in Chomsky and Halle (1968), Kiparsky (1982). There are, however, cross-linguistic differences in degree of the phonetic variation (Öhman 1966, Ladefoged 1967, Clumeck 1976, Lubker and Gay 1982, Magen 1984, Manuel and Krakow 1984, Cohn 1988, Huffman 1988, Manuel 1990, Choi and Keating 1991, among others). For example, the degree of fronting of the velar consonant in [ki] is much larger in French than in Spanish or Dutch (Gussenhoven and Jacobs 1998). Is this phonetic difference just arbitrary then? Manuel's Perceptual Output Constraints, for example, argues that the number of phonological contrasts in a language determines how much a phoneme can deviate from its target position. Languages with smaller inventory allow more contextual deviation than languages with larger inventory with less possibility of

---

* Dept. of English Language and Literature, Ewha Womans University.

confusing perceptual distinctiveness among phonemes (e.g., Manuel and Krakow 1984; Manuel 1990). The purpose of this paper is to provide empirical evidence for the phonetic variation of segments, focused on the CONSONANTAL variation due to contextual vowels, which is explainable by the functional principles, ease of articulation and maintenance of contrasts.

Phonological assimilation is the process that transfers particular articulation to an adjacent sound. For the speaker, there are obvious advantages in extending the particular articulation, since it enhances the economy of speech production. There is also a tendency for languages to avoid repetition of the same phonological element, e.g., dissimilation, the phenomena related to the Obligatory Contour Principle. For the listener, it would be better to hear contrasts by enhancing distinctiveness between phonological elements (e.g., Gussenhoven and Jacobs 1998).

As for the phonetic variation, Flemming (1997) provides preliminary evidence that languages with more high vowels (Finnish, German) show smaller degree of vocalic deviation from their targets than languages with less high vowels (English, Farsi). In this paper, I will show that languages with more high vowels (French with [i, y, u]) complementarily allow *larger consonantal deviation* from its target than languages with less high vowels (English with [i, u]). This will be interpreted as achieving the economy of articulation to a certain extent in order to avoid otherwise extreme articulatory movement to be made in CV syllables. The second study shows that the Russian plain bilabial consonant allows less amount of undershoot than does the English bilabial consonant. This is probably due to the stricter demand on maintaining the consonantal contrasts, plain vs. palatalized, existing only in Russian.

2. Coronal consonants in the context of back vowels

The back vowels have the low second formant (F2) frequency, which is generally raised in the context of the coronal consonants with high F2 targets. In other words, as shown in Figure 1, the back vowel [u] in *who* shows relatively low F2, but the same vowel in *do* shows higher F2 because of the influence of the adjacent coronal consonant with high F2.

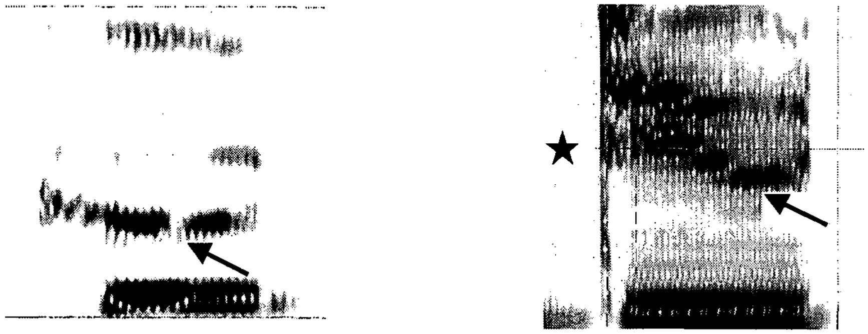

Figure 1. Sample spectrograms of *who* and *do*.

$F2_{T(arget)}$ is the estimated target value of a vowel, e.g., the vowel F2 in *who*, assuming that the laryngeal context have only the least amount of influence on vowel formants (Flemming 1997). The $F2_{V(owel)}$ is the F2 frequency at the steady state of a vowel, e.g., the vowel F2 in *do*. The amount of undershoot (U) is calculated by taking differences between the $F2_T$ and the $F2_V$ (Flemming 1997). $F2_{C(onsonant)}$ is the F2 frequency at the release of a consonant, which is star-marked on the spectrogram in Figure 1.

Five English (Group E) and five French (Group F) speakers read the target words produced in carrier phrases five times each.

(1) English: Say "who" to me.
Say "two" to me.
French: Dites-moi "ou" deux fois.
Dites-moi "tous" deux fois.

The results are presented in Table 1 below. It was confirmed that the degrees of vowel undershoot were larger in English (503 Hz) than in French (189 Hz), and the difference by language was statistically significant (cf. Flemming 1997).

Table 1. Group E and Group F; Repeated measures ANOVAs with speaker nested under group as a random factor.

	Group E	Group F	F Ratio	DF	Prob>F
$F2_T$	1,061 Hz	817 Hz	6.01	1, 32	$p<0.05$
$F2_C$	2,234 Hz	1,760 Hz	55.06	1, 32	$p<0.001$
$F2_V$	1,564 Hz	1,006 Hz	22.71	1, 32	$p<0.01$
U	503 Hz	189 Hz	6.00	1, 32	$p<0.05$

What draws our particular attention here is that the consonant F2s also vary between

English and French. French shows a lower consonant F2 (1,760 Hz) than English does (2,234 Hz), and this difference was statistically significant. Where does this difference come from then? I hypothesize that when a language requires more vowel distinction, the consonant in the same syllable tends to allow more undershoot to avoid otherwise extreme amount of articulatory movement.

In order to examine the degrees of consonantal variation due to the neighboring vowels, I measured English [di] and [du] and French [di], [dy] and [du] produced in carrier phrases (two males and one female in each language). As presented in the sample results in Figure 2, the consonant F2s of English and French systematically display a clear contrast. English shows smaller consonant variation due to the adjacent vowel, while French shows larger consonant variation. As presented in Table 2, I calculated the differences between the consonant F2 in [di] and the consonant F2 in [du]. The mean values were 112 Hz in English and 485 Hz in French. Therefore, there seems to be a clear difference in the amount of consonantal variation.

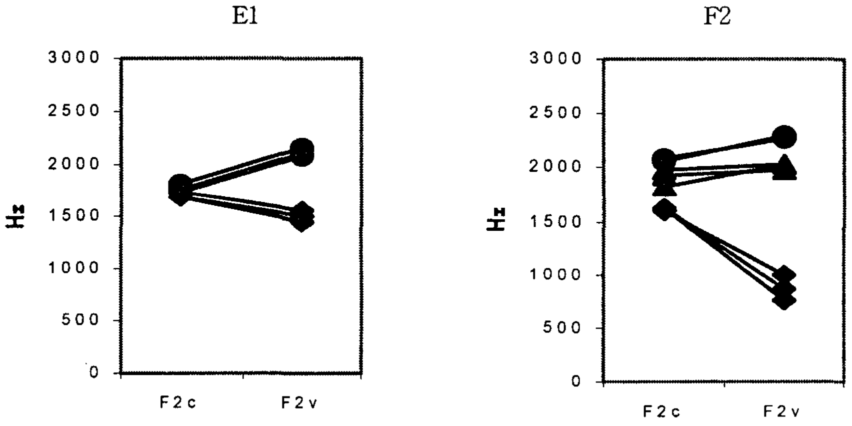

Figure 2. Consonant variation due to adjacent vowels: Circles at the endpoints for [di], diamonds for [du] and triangles for [dy]

Table 2. Mean values of [F2$_C$([di]) - F2$_C$([du])]

E1 (M)	E2 (M)	E3 (F)	F4 (M)	F5 (M)	F6 (F)
60 Hz	67 Hz	210 Hz	453 Hz	304 Hz	699 Hz

(M for male and F for female speech).

However, the story is not complete yet. It is because it may be the case that the French coronal target F2 is inherently low and the large undershoot occurs in the syllable [di], but not in [du]. Therefore, we need to know the absolute consonant target values. A few studies used a formula to estimate the target values·for consonants, called F2$_{L(ocus)}$.

For example, Sussman, McCaffrey, and Matthews (1991) computed locus values for coronals by calculating the point of intersection of the locus equations with the line with slope 1 (i.e. $F2_C = F2_V$). I was able to calculate the loci (i.e. consonant targets) for coronals in English and French from the full locus equations by using the following formula (also Flemming p.c., Klatt 1987, Crowther 1994).

(2) Where $F2_C = F2_V$ for the locus equations $F2_C = mF2_V + c$,
$F2_C = (F2_C - c) / m$
$F2_C = (1 - m) / c$ = locus
(*m* and *c* are the slope and y-intercept, respectively, of the locus equation.)

The locus equation (to be discussed again in the next section) is the regression fit of the consonant F2s in relation to their coarticulated vowel F2s in CV syllables (originally found by Lindblom 1963). To get the locus equations for the coronal stops, first, we recorded the dV syllables, V being all the vowels in each language, as presented in (3) below. Then, we measured $F2_C$ and $F2_V$ of all the syllables. and plotted the $F2_V$'s on the X-axis and the $F2_C$'s on the y-axis. The sample locus equations are presented in Figure 3. Table 3 presents individual data.

(3) English - Say "dVt" to me.
["dVt" = deet, dit, debt, dat, dut, dot, doot, daught]
French - Dites-moi "dVt" deux fois.
["dVt" = dite, dette, dêtte, dute, deute, datte, doute, dot]

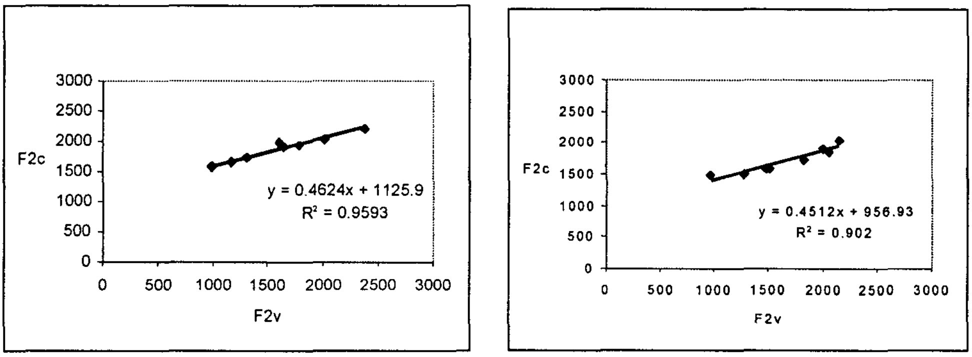

Figure 3. The sample locus equations for the coronal stops in English (left) and French (right). Data points are the mean values of five repetitions.

Table 3. Individual data.

speaker	slope	y-intercept	R^2
E1 (M)	0.58	847	0.93
E2 (M)	0.46	1,126	0.96
E3 (M)	0.33	1,381	0.66
E4 (F)	0.32	1,472	0.83
E5 (F)	0.38	1,442	0.78
F1 (M)	0.25	1,308	0.81
F2 (M)	0.45	957	0.90

With the slope and y-intercept values in Table 3, we could calculate the estimate values of the consonantal targets, using the formula in (2). Table 4 shows the results.

Table 4. The loci for coronal consonants ($F2_L$) in English and French and the undershoot values of the coronal consonants in the context of back vowels (U_C): in Hz.

	E1	E2	E3	E4	E5	F1	F2
$F2_L$	2024	2093	2074	2175	2323	1739	1740
U_C	170	121	58	44	37	162	255

As shown in the second row of the Table 4, the loci are remarkably consistent within English and French and quite different between them. For the male speakers, the loci were around 2050 Hz in English and around 1740 Hz in French. In other words, there *is* a difference in targets.

We also wanted to know whether there is also a difference in the degree of consonant undershoot. Undershoot values of a consonant (U_C) were calculated by taking the difference between the estimated consonant target F2 ($F2_L$) and an F2 of the consonant in a specific vowel context ($F2_C$).

(4) $U_C = F2_L - F2_C(do)$

As presented in Table 4 and Figure 4, even though French coronals inherently show lower F2 than English, they also show larger amount of undershoot in the context of back vowels than English coronals do.

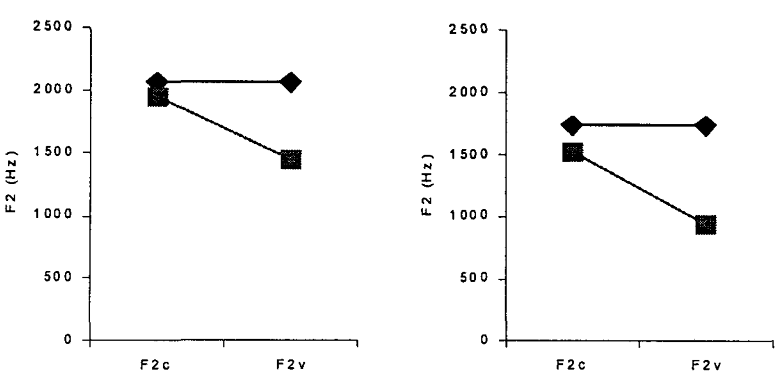

Figure 4. Undershoot of coronal consonants in the back vowel contexts: Mean $F2_L$, $F2_C$, and $F2_V$ values in English (left) and French (right). The horizontal line in each language indicates the mean $F2_L$ value.

To conclude, while English shows a larger degree of vowel variation due to adjacent consonant, French shows a large degree of consonant variation due to the adjacent vowel. A possible explanation regarding this contrast is that with high demand on the achievement of the vowel values French has lower $F2_L$ and allows larger consonant undershoot in order to avoid extreme articulatory movement to be made in coronal consonant and back vowel sequences.

3. Bilabial consonants in vowel context

As is well-known, Russian has the complex consonant inventory. The complexity is mainly due to the existence of the full set of plain and palatalized consonants for each stop place of articulation. For example, the Russian words for 'to take' [brat^j] and 'brother' [brat] are distinguished only by palatalized and non-palatalized coronal consonant. In English, for example, during the production of bilabial stops, tongue body can anticipate the articulation of the following vowel, resulting in a large degree of coarticulation. In Russian, however, a large amount of consonantal undershoot would endanger its contrast with its corresponding palatalized consonant (Öhman 1966; Cohn 1988). The second experiment compares the overall degree of undershoot of the plain bilabial consonants between English and Russian.

As an index of the overall degree of consonantal undershoot, the *locus equation slope value* was employed. As briefly discussed in the last section, plotting the contextual $F2_V$ on the X-axis and the corresponding $F2_C$ on the Y-axis, the function is strikingly linear, called the locus equation (Lindblom 1963). The reason why the slope value of the locus equation is a reasonable index of consonantal coarticulation is as follows. As Krull (1988) discusses, a steep slope is associated with large coarticulation, because it indicates that $F2_C$ changes more with vowel context. A shallow slope is associated with small coarticulation, since $F2_C$ remains fairly stable with adjacent vowels.

The hypothesis to be tested is then that the locus equation slopes are shallower in Russian, indicating less vowel influence on adjacent stops, than in English. Speech materials used in each language are as follows.

(5) English - Say "bVt" to me.
["bVt" = beat, bit, bet, bat, but, bot, boot, bought]
Russian - Povtori "bVd" eshche raz.
'Repeat "bVd" one more time.'
["bVd" = bed, bad, bud, bod]

The experimental results are presented in Figure 5 and Table 5. The locus equation slopes for the English bilabial consonants was steep, the mean value of four speakers being 0.85 (range 0.84 to 0.86) with a low y-intercept (mean=185.99 Hz, range 123 to 221), and characterized by a tight clustering of points around the regression line (mean R_2=0.97). The locus equation slopes for the Russian bilabial stop were shallower than those of English, the mean of four speakers being 0.53 (range 0.45-0.57) with a higher y-intercept (mean=406.35 Hz, range 247-630). Data points were also tightly clustered around the regression line (mean R_2=0.95). Therefore, comparison of slope for the English and Russian bilabial consonants revealed flatter slopes for Russian than English. The difference in slope by language was statistically significant (by ANOVAs with speaker nested under group as a random factor, [$F(1,7)=139$; $p<0.0001$]).

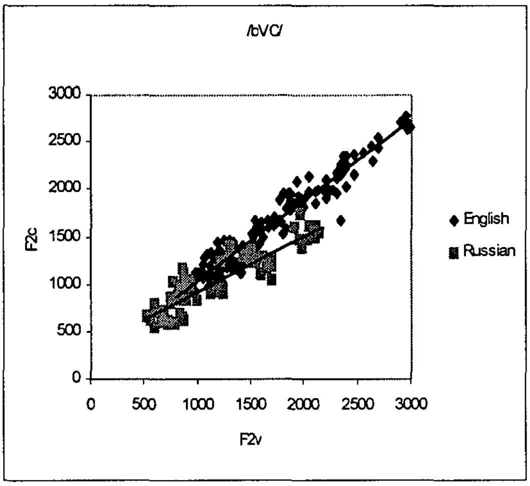

Figure 5. Raw data points generated by four English and four Russian speakers.

Table 5. Comparisons of the locus equation parameters (mean values).

	Slope	y-intercept
English	0.85	186
Russian	0.53	406

The basis for this difference in slopes for the bilabial tokens can be seen most clearly in the spectrograms of Russian *bed* and English *bet* in Figure 6. The Russian bilabial stop shows minimal undershoot, pointing closely to a target position, but the English bilabial stop shows maximal undershoot with F2 at release being almost equal to F2 in the following vowel.

This can also be seen in Figure 7, which graphs the mean values of the native consonant F2s and vowel F2s in each syllable type. Although the difference in the transition rates between Russian *bad* and English *but* (with similar vowel F2 values) was less striking, the Russian bilabial again showed less consonant undershoot. Note that the Russian *bad* lies on a line with Russian *bed* and off the English line, as observed in divergence of English and Russian lines in range 1400-1700 Hz for F2v.

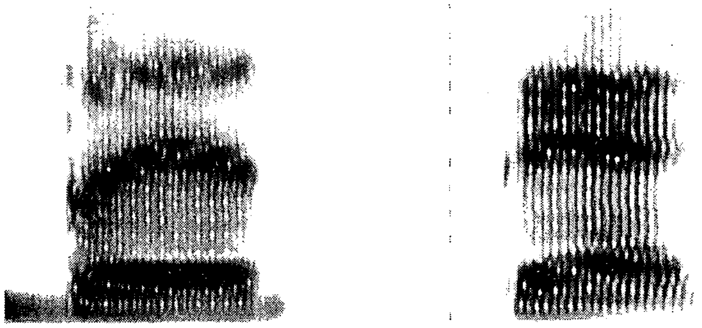

Figure 6. Sample spectrograms of Russian *bed* (left) and English *bet* (right).

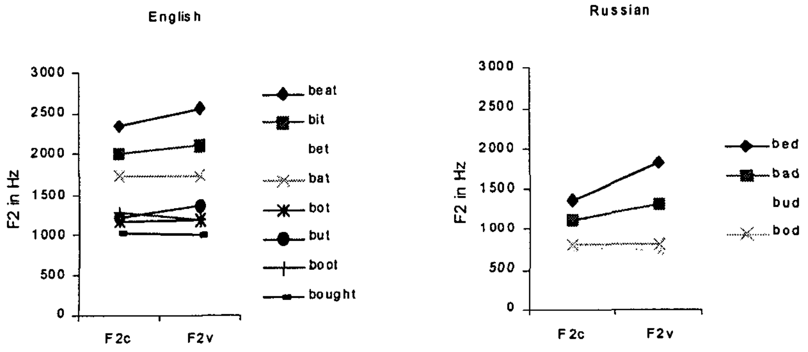

Figure 7. Plots of mean values of the F2c's and F2v's.

4. Conclusion

This paper presented empirical evidence that the gradient assimilation of a segment towards the neighboring sounds, which has often been attributed to universal phonetic implementation, is in fact language-specific (the references cited in section 1). The two case studies were examined. French, with larger demand on maintaining contrasts among high vowels, complementarily allows larger consonantal undershoot than English does, in order to reduce the articulatory movement in the CV syllables to a certain extent. The Russian plain bilabial consonant allows less amount of undershoot than does the English consonant, probably because of the stricter demand on maintaining consonantal contrasts in Russian. These examples demonstrate clear indications that the functional principles, ease of articulation and maintenance of contrasts, account not only for the categorical, phonological processes, but also for the gradient, phonetic variation (Lindblom 1983), and that (at least part of) phonetic phenomena are not simply automatic or arbitrary, but are with 'good' reason.

Acknowledgement

I would like to thank Edward Flemming and two anonymous reviewers of *Korean Journal of Speech Sciences* for helpful suggestions and comments. An earlier version of this paper was presented at Seoul International Conference on English Linguistics 2001. I am grateful to the audience for useful questions and comments.

References

Choi, J. D. & P. Keating. 1991. "Vowel-to-vowel coarticulation in three Slavic languages." *UCLA Working Papers in Phonetics,* 78, 78-86.

Chomsky, N. & M. Halle. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Clumeck, H. 1976. "Patterns of soft palate movements in six languages." *Journal of Phonetics,* 4, 337-351.

Crowther, C. S. 1994. "Modeling coarticulation and place of articulation using locus equations." *UCLA Working Papers in Phonetics*, 127-148.

Cohn, A. 1988. "Quantitative characterization of degree of coarticulation in CV tokens." *UCLA Working Papers in Phonetics*, 51-59.

Flemming, E. S. 1997. "Phonetic optimization: Compromise in speech production." In *University of Maryland Working Papers in Linguistics 5: Selected phonology papers from H-OT-97.*

Gussenhoven, C. & H. Jacobs. 1998. *Understanding Phonology.* Arnold.

Huffman, M. K. 1988. "Timing of contextual nasalization in two languages." *UCLA Working Papers in Phonetics*, 68-76.

Kiparsky, P. 1982. "Lexical phonology and morphology." In *Linguistics in the morning calm*, ed. by I. S. Yang, 3-91. Seoul: Hanshin.

Klatt, D. H. 1987. "Review of text-to-speech conversion for English." *Journal of the Acoustical Society of America*, 82, 737-793.

Krull, D. 1987. "Second formant locus patterns as a measure of consonant-vowel coarticulation." *PERILUS* V, 43-61. University of Stockholm.

Ladefoged, P. 1967. "Linguistic phonetics." *UCLA Working Papers in Phonetics* 6.

Lindblom, B. 1963. "Spectrographic study of vowel reduction." *Journal of the Acoustical Society of America*, 35, 1773-1781.

Lindblom, B. 1983. "Economy of speech gestures." In *The Production of Speech*, MacNeilage, P. (ed.). New York: Springer-Verlag.

Lubker, J. & T. Gay. 1982. "Anticipatory labial coarticulation: experimental, biological and linguistic variables." *Journal of the Acoustical Society of America*, 71, 437-448.

Magen, H. S. 1984. "Vowel-to-vowel coarticulation in English and Japanese." *Journal of the Acoustical Society of America*, 75, S42.

Manuel, S. Y. & R. A. Krakow. 1984. "Universal and language particular aspects of vowel-to-vowel coarticulation." *Haskins Laboratories Status Report in Speech Research*, SR-77/78, 69-78.

Manuel, S. Y. 1990. "The role of contrast in limiting vowel-to-vowel coarticulation in different languages." *Journal of the Acoustical Society of America*, 88, 1286-1298.

Öhman, S. E. G. 1966. "Coarticulation in VCV utterances: Spectrographic measurements." *Journal of the Acoustical Society of America*, 39, 151-168.

Sussman, H. M., H. A. McCaffrey. & S. A. Matthews. 1991. "An investigation of locus equations as a source of relational invariance for stop place categorization." *Journal of the Acoustical Society of America*, 90, 1309-1325.

Received: July 25, 2001.
Accepted: September 4, 2001.

▲ Eunjin Oh
Department of English Language and Literature
Ewha Womans University
11-1 Daehyun-dong, Seodaemun-gu, Seoul 120-750, Korea
Tel: +82-2-312-7605
E-mail: ejoh-7@hanmail.net