The role of hypercorrection in the acquisition of L2 phonemic contrasts

Fred R Eckman, Gregory K Iverson, and Jae Yung Song
University of Wisconsin-Milwaukee, USA

Abstract
This article reports empirical findings from an ongoing investigation into the acquisition of second-language (L2) phonemic contrasts. Specifically, we consider the status and role of the phenomenon of hypercorrection in the various stages through which L2 learners develop and internalize a target language (TL) contrast. We adopt the prevailing view in both sociolinguistics and second language acquisition studies that hypercorrection results from a certain amount of linguistic insecurity on the part of the speaker. Based on 53 Korean speakers' production of English target phonemes, we conclude that a series of hypercorrection errors may well represent the final stage in the acquisition of a contrast, and further, that in order for hypercorrection to occur, there must be a formal connection between the TL contrast being acquired and the phonological structure of the learner's native language.

I   Introduction

'Hypercorrection' is a technical term that has been employed extensively in studies of language variation and linguistic change to describe the extension of a linguistic form beyond its regular usage, sometimes resulting in an erroneous production. As characterized by Labov (1966, 1972) and employed in sociolinguistics (e.g. Wolfram, 1991) as well as in historical linguistics (e.g. Campbell, 1998), hypercorrection involves awareness on the part of speakers of language varieties that are associated with differing

Corresponding author:
Fred R Eckman, Department of Linguistics, University of Wisconsin-Milwaukee, 3243 N. Downer Avenue, Curtin Hall, Miwaukee, WI 53211, USA.
Email: eckman@uwm.edu

degrees of prestige. The rendition of a form by a speaker of a less prestigious variety in an attempt to have it match a more prestigious pattern, but which in the process overshoots the mark and thereby results in an 'incorrect' form, is thus an instance of hypercorrection. This concept has also played a role in the characterization of second language (L2) errors, as in, for example, Janda and Auger's (1992) study on the pronunciation of English word-initial [h] by native speakers of French, a language in which neither [h] nor aspiration occurs.

In this article we report data gathered to test two hypotheses about the nature and role of hypercorrection in the acquisition of an L2 phonemic contrast. The general thrust of the proposal is that hypercorrection constitutes a near-final, if not the final, state of acquisition in certain well-defined cases. In particular, we identify the kinds of interlanguage (IL) changes that characterize a learner's progress in the acquisition of novel phonemic contrasts, changes, which in turn lay the phonological foundation on which the hypercorrection is based.

The article is structured as follows. In Section II, we set the context for the study by outlining previous discussions in the literature on hypercorrection, with particular emphasis on second language acquisition (SLA). We review the facts reported by Janda and Auger (1992), in their study of native speakers of French acquiring English /h/. We then briefly recapitulate the findings from one of our recent studies describing the systematic occurrence of hypercorrection errors. We conclude this section by proposing two hypotheses about the phonological mechanism underlying hypercorrection errors in L2 pronunciation, noting the empirical implications and the kinds of data that would test our claims. Section III describes the gathering and scoring of the production data elicited to test the hypothesis, and Section IV reports the findings and evaluates their implications for SLA.

II Background

1 *Characterization of hypercorrection*

In seminal work on hypercorrection in sociolinguistics, Labov (1963, 1966) distinguished the phenomenon into two types: statistical and structural. Statistical hypercorrection, also referred to as 'quantitative hypercorrection' (Janda and Auger, 1992, and numerous works cited therein), results from a stylistic shift in which a speaker of one variety emulates a more prestigious version of the language, and overuses a particular form, thereby exceeding the frequency norms of the targeted variety. The classic example provided by Labov's work is the variable occurrence in New York City social dialects of post-vocalic /r/ in words such as *guard* and *floor*. Labov found that the incidence of post-vocalic /r/ varied systematically across socio-economic groups according to style of speech. All of the groups increased the use of post-vocalic /r/, a prestige variable in New York City, as the speech style became more formal. The interesting finding was that, in the two most formal styles of speech, the lower middle class pronounced post-vocalic /r/ more frequently than did the upper middle class, which otherwise, as the social class with highest prestige, showed the greatest use of this variable. This form of hypercorrection is statistical in that speakers of one variety, in emulating a more prestigious variety, ended

up producing the feature in question to a greater extent in formal speech than did the variety they were emulating.

The other kind of hypercorrection Labov identified is structural hypercorrection (labeled 'qualitative hypercorrection' in Janda and Auger, 1992), which arises when speakers extend a prestigious linguistic property to environments in which it otherwise does not occur. A well-known example involves speakers of Cockney English, which ordinarily lacks the sound [h], thus producing words such as *house* as [aws] or [ʔaws]. Realizing that this *h*-less pattern of speech does not match the prestige standard, Cockney speakers may change their pronunciation of some words by adding [h] at the beginning, even when these words in the prestigious variety of English begin with a vowel. This strategy, for example, causes Cockney speakers to introduce a gratuitous [h] in, and so hypercorrect, the word *out* as [hawt].

Hypercorrection has also been reported in studies on language change, under the rubric of analogy. Campbell (1998) lists hypercorrection as an example of analogical change resulting from speaker attempts to change a form in a less prestigious variety in order to make it conform to how it is rendered in a more prestigious variety. Speakers sometimes overproduce the form, resulting in erroneous outcomes with respect to the variety that is being mimicked, as in the introduction of /r/ phrase-finally (*this is a good idea*r) by speakers imitating the otherwise post-vocalic *r*-less dialect which has 'linking-*r*', but only between vowels (*this idea*r *is good*, versus *this is a good idea*). What unifies all of these types of hypercorrection is that the speakers are aware of the variation that surrounds a given linguistic feature, resulting in the overproduction of that feature in an effort to emulate a particular, prestige-bearing variety of speech. Thus, within the context of sociolinguistics, Wolfram (1991: 155) states straightforwardly that 'The source of hypercorrection is *linguistic insecurity*.' Likewise, Knowles (1978: 86) characterizes hypercorrection in terms of a speaker being 'sufficiently concerned about his speech to aim at the standard pronunciation'. In fact, in Labov's (1963, 1966) studies of New York /r/, the mere existence of qualitative hypercorrection is taken to be the indicator of the speakers' linguistic insecurity.

2   Hypercorrection in SLA

Structural hypercorrection has also been identified in studies on SLA. Janda and Auger (1992), in particular, provide an extensive and thorough discussion of hypercorrection as it has been investigated by both historical linguists and sociolinguists over the past century. Also within the context of SLA, Odlin (1989: 38) categorizes hypercorrection as an instance of crosslinguistic influence in which the learner overreacts to a particular influence from the native language (NL). And Siegel (2003: 200) equates hypercorrection as used in sociolinguistics with overgeneralization in SLA.

The proposals that we will argue for in this article impinge on the three claims above that have been made about hypercorrection, both in general and specifically with respect to SLA. The first is Odlin's statement that hypercorrection errors by L2 learners stem from crosslinguistic influence. The second is Siegel's assertion that hypercorrection in SLA is equivalent to overgeneralization. And the third is the commonplace view that speakers produce hypercorrection errors as they emulate a prestige language variety. The

issue that must be addressed in the discussion of hypercorrection in SLA is the argument for considering the target language (TL) to be prestigious. We take up each of these topics in turn, beginning with Odlin's assertion about crosslinguistic influence.

One of the points we seek to establish in this article is that hypercorrection errors are different from straightforward crosslinguistic influence. As the term is generally used in the SLA literature, crosslinguistic influence on L2 learner pronunciation labels a pattern of errors brought about by the incorrect application, or extension, of NL patterns to the pronunciation of the interlanguage. This type of error pattern has long been recognized in the SLA literature as non-first-language-influenced substitution, though not as much has been written about hypercorrection errors (e.g. Kohler, 1971) per se. Thus, for example, Nemser (1971) documented non-first-language-influenced substitution error types in a study of native-speaking Hungarians learning English, and Johansson (1973) attested non-transfer errors in the acquisition of L2 Swedish. In a similar vein, Hecht and Mulford (1982) studied the interaction of transfer and developmental errors. And, finally, non-first-language-influenced errors also form the crux of the Ontogeny Phylogeny Model (Major, 2001).

In the case of hypercorrection errors, however, L2 learners are not following the patterns of their NL; in fact, if the learners did apply their NL patterns, the errors in question would not occur. This is because hypercorrection errors result from learners producing forms that are actually counter to the NL patterns. For example, Janda and Auger (1992) report that their French-speaking learners of English incorrectly produce initial [h] in words that begin with a vowel in English. The segment [h] is absent in French, of course, so the production of an *h*-initial English word like *head* as [ɛd] by a native speaker of French can be attributed to crosslinguistic influence. However, such a speaker's production of English *ache* as [hek] is not a case of direct crosslinguistic influence, since, if the NL pattern were followed, there would be no initial [h], and the form would be target-like. Following Janda and Auger, we classify forms such as [hek] as hypercorrection errors, not direct language transfer errors. We will expand on this point below.

Within sociolinguistics, we believe Siegel's (2003: 200) characterization of hypercorrection errors in SLA as simply overgeneralizations also misses the mark. There are numerous reported examples of L2 learners overgeneralizing patterns in the TL such that the errors do not parallel hypercorrections in sociolinguistics. For example, a speaker's regularization of the past tense of English irregular verbs so as to result in *runned* instead of *ran* is a case of overgeneralization, but cannot be interpreted as a hypercorrection error inasmuch as *runned* is not found in the prestige standard variety of the language. Thus, the concept of overgeneralization, as it is normally understood, does not adequately characterize the notion of hypercorrection in SLA, precisely because overgeneralization inappropriately extends a TL pattern (*runned, via categorization of *run* as a weak verb), whereas hypercorrection involves not only going counter to an NL pattern (*[hek] for *ache*, via suppression of the French prohibition of [h]), but also incorrectly implementing a newly learned structure or sound.

We now turn to the third issue to be discussed, the rationale for L2 learners' considering the TL to be prestigious. Here, we rely on two studies of hypercorrection in SLA, Janda and Auger (1992), mentioned above, and John and Cardoso (2009), which also deals with the L2 acquisition of English /h/ by native speakers of French. Janda and

Auger make the point that the term 'hypercorrection' is appropriate for L2 learners in this case because the TL, English, is the prestige variety that the L2 learners are emulating, not in absolute terms such that English is considered to be more prestigious than French, but rather in the sense that the TL is the variety to be emulated if learners are to speak that language intelligibly. John and Cardoso (2009) come to essentially the same conclusion regarding the prestige of the TL, although they disagree with some of the specifics of Janda and Auger. John and Cardoso suggest that a more likely reason for the value accorded the phoneme /h/ by francophones is, first, that it is more prestigious to speak English like a native speaker, and, second, that it is stigmatized not to articulate [h], since *h*-deletion is one of the most salient and least favored features of francophone speech. In short, the authors of both studies agree that, in an L2 learning context, the TL is considered to be more prestigious than the NL, even if not absolutely so.

The proposal we advance here is that, with respect to the L2 acquisition of phonemic contrasts, hypercorrection represents a late stage of acquisition for some kinds of phonemic contrasts, depending on the structure of the NL phonology relative to the TL contrast to be learned. That hypercorrection errors systematically occur later, rather than earlier, in the acquisition process is not a novel idea. This point is made in Janda and Auger, and is consistent also with Major's (2001) Ontogeny Phylogeny Model. The late occurrence of such errors also follows automatically from our characterization of hypercorrection as the incorrect extension of a newly learned contrast. That is, because of the relative prestige of the TL, hypercorrection errors are always in the direction of the most recently acquired member of the contrast.

As to the underlying cause of hypercorrection errors, we endorse the mainstream view identified above that the phenomenon arises out of the 'linguistic insecurity' that affects speakers who, in emulating the structures of the more prestigious language – the TL – are attempting to avoid previously recognized errors. Though we have nothing to add to this understanding, the goal we set here is to lay out the mechanism whereby hypercorrection errors occur once L2 learners have reached the point at which they find themselves in a position of linguistic insecurity about certain TL pronunciations.

3 Three contrasts under consideration

We now turn to the specifics of our study. The three combinations of NL and TL phonological structure that are relevant to this investigation are listed in (1) below. We will argue that the language contact situations represented in (1b) and (1c) involve a stage of hypercorrection errors, whereas that in (1a) does not.

(1) a. The NL lacks analogs of the two TL sounds in contrast.

   b. The NL has only an equivalent of one of the two contrasting TL sounds.

   c. The NL has both of the TL sounds, but as allophones of the same phoneme in complementary distribution.

We exemplify with the case of (1c) above with Koreans acquiring the English /s/–/š/ contrast, as 'lax' [s] and [š] are allophones of the same phoneme in the NL (paralleling laryngeally 'tense' [s'] and [š'], which are allophones of a separate fricative). The

distribution of these segments can be characterized informally by a rule such as (2), which specifies that a sibilant is realized as palatoalveolar before high front vocoids (/i/ and /j/), i.e. as [š], [š'] in the case of fricatives, [č], [č'], [čʰ] in the case of affricates. Elsewhere in the native vocabulary, fricatives are alveolar ([s], [s']) while affricates have moderately postalveolar release ([c], [c'], [cʰ] = IPA [t͡ɕ], [t͡ɕ'], [t͡ɕʰ]) (cf. Eckman and Iverson, 2013; Kim, 2001).

    (2)   sibilant → palatoalveolar / __ high front vocoid

Transfer of this rule into a learner's IL grammar causes errors on TL words in which [s] occurs before high front vowels, such as *see* or *sip*, rendering these as *[ši] and *[šɪp]. Therefore, the task of the Korean learner of English is to suppress the application of rule (2) in order to permit sibilant phonemes without palatal articulation to occur before /i/. Failure to suppress (2) represents an early stage of learning in which NL transfer determines IL pronunciation, whereas successful suppression represents a late stage in which TL-like pronunciations have been achieved. This progression is itself staged, however, with effects of the NL rule persisting in derived or crucially intermorphemic contexts (so that *messing → meshing*) even as the rule is suppressed morpheme-internally (*see ≠ she*), as we have detailed elsewhere (most recently, Eckman and Iverson, 2013). Acquisition of a TL contrast of the type in (1c) is thus gradual, with a phonemic distinction established in some contexts before it is in others: preferentially in words for which the context-sensitive NL rule is not relevant (e.g. *sew* versus *show*) and morpheme-internally in words for which it is (*see* versus *she*). Ultimately, the NL rule comes to be suppressed intermorphemically as well, so that the TL phonemic contrast is in evidence throughout (*messing ≠ meshing*).

    The language contact situation in (1b) is exemplified by a native-speaker of Korean attempting to learn the English /p/–/f/ contrast, and by French native speakers learning to produce the English contrast between syllable-initial /h/ and onsetless syllables. French lacks [h], though it has onsetless syllables; Korean has [p], but not [f]. Labial true consonants in Korean must be stops, either the nasal /m/ or one of the three contrasting plosives /p/, /p'/, /pʰ/; in particular, there is no labial fricative, either as a phoneme in its own right (/f/) or as an allophone ([f]) of one of the stops. These facts can be characterized formally by a context-free rule or constraint in the grammar of Korean to the effect given in (3). If transferred into the IL by Korean-speaking learners of English, successful or target-like pronunciation requires suppression of this constraint, which militates against fricatives at the labial place of articulation irrespective of context:

    (3)   Labial → stop

Persistence of the constraint in (3), that is, accounts for early IL errors such as the pronunciation of *fork* with an initial aspirated [pʰ], whereas later suppression of (3) enables the TL-correct pronunciation with initial [f].

    Finally, the language contact situation in (1a) can be illustrated by Korean L2 learners of English acquiring the /f/–/v/ contrast, two sounds which, consistent with (3), do not occur in Korean. As these sounds are absent from Korean, there is no NL phonological rule or constraint that relates them and which an L2 learner can transfer into the IL grammar. Therefore, according to the hypotheses developed in (5) below, there should be no systematic hypercorrection errors between these two sounds.

We now turn to the facts concerning hypercorrection errors in L2 phonology by focusing on two earlier studies, Janda and Auger (1992) and Eckman and Iverson (2013). Janda and Auger (1992) studied NL transfer and hypercorrection errors by six native speakers of French learning English. Productions were elicited across five tasks: free conversation, reading a connected text, reading isolated sentences, reading a word list, and reading a set of minimal pairs. An example of an NL transfer error in this study was the deletion of [h], as in the pronunciation of *head* as [ɛd], and an instance of a hypercorrection error was producing the word *ache* as [hek]. Although the results varied by participant, and according to task, the authors were able to cite a few patterns. What is noteworthy is that all of the participants produced some NL transfer errors in at least one of the tasks, and all of the participants produced some hypercorrection errors in at least one of the tasks. Moreover, the percentage of hypercorrection errors aggregated across tasks was small, ranging from a low of 0.08% to a high of 1.9%.

The purpose of our earlier study by Eckman and Iverson (2013) was to track the acquisition of the English /s/–/š/ contrast as a function of two phonological contexts, in which general principles of phonology predict that there would be implicationally related patterns of acquisition. Analysis of the data revealed systematic NL transfer and hypercorrection errors by the participants. An NL transfer error is the production of a word like *see* as [ši], as this is an instance of the NL allophonic pattern intruding into the IL. A hypercorrection error is exemplified by a learner pronouncing a word containing [š] before a high front vowel as [s], as in *bushy* being rendered as [bʊsi]. In this case, if the learner had followed the NL patterns producing [š] before [i], the form would have been TL-like.

The results from this study were also variable, but we noted the following patterns. First, 12 of the 26 participants (46%) produced both transfer and hypercorrection errors, with higher percentages of hypercorrection occurring in the productions of participants who demonstrated the contrast according to an 80% threshold for acquisition. Second, the hypercorrection errors occurred only in words containing [š] before a high front vowel. Thus, we attested numerous instances of a word like *shin* being rendered as [sɪn], but found no examples of *shoe* being pronounced as [su].

Based on these empirical studies of hypercorrection in SLA, we can summarize the relevant facts as in (4).

(4) Hypercorrection in SLA

    a. Hypercorrection errors occur later rather than earlier in the acquisition process.

    b. NL transfer errors and hypercorrection errors can co-exist.

The hypotheses advanced in (5) below seek to capture the nature of the mechanism underlying hypercorrection errors in L2 phonology.

(5) a. Hypercorrection errors in L2 phonology will occur only under the condition that at least one of the phonemes in the contrast being acquired exists in the NL.

b. With respect to an environmentally conditioned NL rule that is transferred into the IL, hypercorrection errors will occur only in contexts defined by the environment of the rule.

The rationale for, and fundamental claim of, these hypotheses is that systematic hypercorrection errors occur in IL phonological productions only when the NL–TL contact situation is such that (1) the learner must acquire a new TL contrast where at least one of the members of the distinction being learned occurs also in the NL; or (2) the learner must acquire the production of an NL segment in an environment where it does not occur in the NL. Where no such overlap between the NL and TL exists, the prediction is that there will be no systematic hypercorrection errors.

There are two interesting implications of these claims. The first is that, in the case of Korean native speakers learning the English /s/–/š/ contrast, the hypercorrection errors should occur only in the environment of the transferred NL allophonic rule. Native speakers of Korean, therefore, should produce hypercorrection errors only on words where [š] occurs before a high front vowel, as in *shingle*, but not in words such as *shoe*. Conversely, since there is no context sensitive rule in the NL relating [p] and [f], we predict that hypercorrections involving this contrast would not be restricted to any one environment.

This leads us to the specifics of the present study.

III Methods

1 Research participants and stimuli

We elicited productions on the three contrasts in question, /s/–/š/, /p/–/f/, and /f/–/v/, from 53 native speakers of Korean. We gathered productions on the /s/–/š/ contrast from 26 participants,[1] on the /p/–/f/ contrast from 12 speakers,[2] and on the /f/–/v/ contrast from 15 participants. Each of the participants belonged to only one of the three contrast groups, and all were between the ages of 18 and 36.

The stimuli used to elicit the productions consisted of a set of 90 words for each contrast, 60 of which were targets (listed in Appendices 1, 2 and 3) and 30 of which were fillers.[3] All are existing lexical items in English, and each target word contained the contrasting segments in one of four positions: word initially before a vowel, word finally after a vowel, word medially between vowels in a monomorphemic word, and medially between vowels at the juncture of a morpheme, namely the suffix -*ing*, as in *hopping*, or -*y*, as in *puffy*.

2 Data gathering and scoring

Several custom programs were written in MATLAB for the purposes of eliciting and recording the productions. The program controlling the elicitations displayed on a computer screen a set of pictures, clues and commands (such as 'Wait' or 'Speak') designed to guide the participant and the experimenter in eliciting the word in question. Words were elicited, not by giving their spelling, but by displaying an image depicting the action or object in question for both monomorphemic words (for example, using a

picture of a man puffing smoke from a cigar to elicit the word *puff*) and then displaying the same picture, but with the cue, 'current action', appearing on the screen one-half second after appearance of the image to elicit the word *puffing*. If the participants could not identify the word on the basis of the image, they were provided with additional on-screen definitions and hints. If the participants still could not identify the intended word, they were presented with a recorded model. The stimuli were presented in a pseudo-randomized order in that all monomorphemic forms were elicited before their related morphologically composite forms. The elicitations were recorded directly onto a hard disc drive at the sampling rate of 44.1 kHz. Participants spoke into a head-mounted microphone at a distance of one inch from the lips and produced the set of 90 words twice, both during the same session.

The data were collected at the University of Wisconsin-Milwaukee and then trans-ferred to the Ohio State University via file transfer protocol where they were tran-scribed by assistants who were blind to the hypotheses. Not only did the transcribers not know what the hypotheses were, they also were unaware of the intended target segments. The transcribers listened to the utterances in question and were focused either on a consonant in word-initial, -medial or -final position, or a word-medial con-sonant occurring before the suffix -*ing* or -*y*. The transcriber's task was then to choose a phonetic description of the segment in question from a menu of several choices. The completed transcriptions were then returned to Milwaukee where they were scored.

For each contrast, the participants produced 120 target words with the contrasting segments in four different environments. There were five monomorphemic words with each of the contrasting segments in word-initial and word-medial positions produced twice during the recording session for a total of 40 words (5 words × 2 contrasting seg-ments × 2 environments × 2 elicitations = 40). Each participant produced an additional 80 words, 40 with the contrasting segments in word-final position and 40 with the seg-ments before the -*ing* or -*y* morpheme juncture (10 words × 2 contrasting segments × 2 two environments × 2 elicitations = 80 words). A participant's performance on the pro-ductions had to reach the 80% criterial threshold for both contrasting segments in a given environment in order for the participant's interlanguage grammar to be credited with having the contrast in that environment. If a participant's performance reached the crite-rial threshold on only one of the segments in a given environment, or did not reach crite-rion on either segment, the participant's IL grammar was scored as lacking the contrast in that environment. For example, a participant had to produce [p] in at least four of the five words in which [p] occurred in word-initial position before a vowel, and likewise for [f], in order for the participant's IL to be given credit for having the /p/–/f/ contrast in initial position.

We scored the kinds of errors participants made in the their attempts to produce each of the contrasts. If a participant erred by substituting, for example [š] for [s] before a high front vowel, we labeled this as an NL transfer error, because [š] is the segment that occurs in this environment in the NL. As another example, if a participant erred by pro-nouncing *five* as *[pajv], this was scored as an NL transfer error because [p], but not [f], is the segment that occurs in the participant's native language. And, finally, if the partici-pant erred by incorrectly producing [f] in words containing [p], as in producing *pan* as *[fæn] or by pronouncing *shin* as *[sɪn], we designated the utterance as a hypercorrection error.

IV   Results

In order to set the stage for the presentation of results, we need to make two points. The first concerns how we report findings in terms of both individualized and aggregated data, and the second pertains to considerations involved in the prediction of errors. After review of these two points, we report the data first in terms of how they bear on the hypotheses in (5), and then in light of the observations listed in (4).

The results are presented individually for each participant, though they will be discussed both individually and in terms of groups as necessitated by the claims we are addressing or the questions we are trying to answer. Our hypotheses make claims about the status of a learner's IL grammar, stating, specifically, that it is the presence of a transferred NL rule in the IL that is the basis for hypercorrection errors. IL grammars are mental systems whose placement in time and space is in the mind of individual learners. Therefore, we must test the claims that are about the state of an interlanguage grammar using individualized data, simply because there is no IL grammar of a group of people, at least not one that can be situated in time and space, just as there is no mind of a group. Thus, we therefore present the results below for each individual participant. However, there are other claims or issues that we will discuss, and that can be better addressed by comparing grouped results. For example, in light of the question as to whether there is a connection between a given participant evidencing the contrast at hand and the kinds of errors that participant produces, we will aggregate our data for purposes of presenting the statistical correlations.

The second point has to do with the variability that can occur with the production of L2 learners' errors and the concomitant uncertainties associated with error prediction. As a precursor to the L2 situation, consider an analogy from syntactic description. It is not the goal of a syntactician to forecast the sentences that speakers of a language will produce. Rather, the syntactician's task is to set out the principles governing the form that utterances of a language can take, once the meaning of the intended utterance has been determined. For example, syntacticians do not aim to predict whether a speaker of English will utter the sentence *The cat chased the mouse*. But once it has been decided that a speaker wishes to express the meaning embodied in this sentence, the task of the syntactician is to specify the grammatical possibilities for doing so, such as *The cat chased the mouse, The mouse was chased by the cat, The mouse, the cat chased (it)*, and so on. Similarly, we do not set as our goal to forecast or prognosticate whether or when L2 learners will make errors, but rather – if they are going to make an error in the pronunciation of a given TL word – to identify the principles governing the errors that the learners make. Proposals are falsified if the errors made fall outside of what is prescribed by the principles involved, and are supported to the extent that the errors conform to what the principles allow; proposals are not falsified if some learners fail to produce certain kinds of errors.

Bearing this in mind, we consider our participants' performance on each of the three contrasts in question with respect to the hypotheses in (5), and in relation to the error types they produced, i.e. whether the participants produced NL transfer errors, hypercorrection errors, or both. The hypotheses are supported if the data summarized in the tables show the existence of two distinct patterns. The first of these is that hypercorrection errors occur systematically in the participants' production of the /s/–/š/ and the /p/–/f/

Table 1. Participants' production expressed as percentages on [s] and [š] before a high front vowel word initially and intermorphemically along with numbers and percentages of NL transfer and hypercorrection errors.

Participants	Environments											
	Word-initial (__ i)						Intermorphemic (__ + i)					
	[s] %	[š] %	HC Errors		NL Errors		[s] %	[š]%	HC Errors		NL Errors	
			n	%	*n*	%			*n*	%	*n*	%
2040	0	100	0	0	10	36	10	100	0	0	18	64
2021	0	100	0	0	10	34	5	100	0	0	19	66
2034	10	100	0	0	9	36	20	100	0	0	16	64
2033	30	100	0	0	7	41	15	100	0	0	17	59
2031	50	100	0	0	5	63	85	100	0	0	3	37
2036	80	100	0	0	2	29	75	100	0	0	5	71
2062	90	100	0	0	1	33	90	100	0	0	2	67
2064	90	100	0	0	1	8	45	100	0	0	11	92
2038	10	100	0	0	9	30	5	90	2	6	19	64
2032	40	100	0	0	6	46	70	95	1	8	6	46
2037	50	90	0	0	6	27	20	100	0	0	16	73
2025	70	70	0	0	6	33	65	75	5	28	7	39
2035	80	100	0	0	2	16	75	75	5	42	5	42
2059	80	100	0	0	2	17	80	70	6	50	4	33
2026	90	100	0	0	1	14	95	75	5	72	1	14
2039	100	70	3	16	0	0	60	60	8	42	8	42
2047	100	100	0	0	0	0	95	95	1	50	1	50
2049	100	90	1	33	0	0	95	95	1	33	1	33
2058	100	100	0	0	0	0	25	100	0	0	15	100
2063	100	100	0	0	0	0	95	95	1	50	1	50
2051	100	100	0	0	0	0	100	95	1	100	0	0
2022	100	90	1	25	0	0	100	85	3	75	0	0
2027	100	80	2	22	0	0	100	65	7	78	0	0
2046	100	100	0	0	0	0	100	90	2	100	0	0
2048	100	90	1	9	0	0	100	50	10	91	0	0
2050	100	100	0	0	0	0	100	100	0	0	0	0

contrasts, but not in the pronunciation of the /f/–/v/ contrast. The second is that, for the /s/–/š/ and the /p/–/f/ distinctions, there should be a general increase in hypercorrection errors as the learners evidence increasing mastery of the contrast in question.

These results are presented in Tables 1 through 3. The tables summarize the relationship among three aspects of the productions of these contrasts. Columns two through six and eight through thirteen show, respectively, a participant's performance on [s] and [š] in word-initial position and syllable-initially across a morpheme boundary before the suffixes -*ing* and -*y*. A participant's production on the two sounds in question is expressed as a percentage of target-like performance. In columns four through seven and 10 through

Table 2. Participants' production expressed as percentages on TL [p] and [f] before a high front vowel word initially and intermorphemically along with numbers and percentages of NL transfer and hypercorrection errors.

Participants	Environments										
	Word-initial (__ i)						Intermorphemic (__ + i)				
	[p] %	[f] %	HC Errors		NL Errors		[p] %	[f] %	HC Errors		NL Errors	
			n	%	n	%			n	%	n	%
2052	100	100	0	0	0	0	95	25	1	6	15	94
2024	100	100	0	0	0	0	100	50	0	0	10	100
2053	100	100	0	0	0	0	100	95	0	0	1	100
2043	100	30	0	0	7	29	100	15	0	0	17	71
2044	100	90	0	0	1	20	85	95	3	60	1	20
2041	90	100	0	0	1	100	100	100	0	0	0	0
2042	100	100	0	0	0	0	85	80	2	33	4	67
2045	100	100	0	0	0	0	95	85	1	25	3	75
2054	100	90	0	0	1	7	90	40	1	7	13	87
2055	100	100	0	0	0	0	20	75	15	71	6	29
2057	100	100	0	0	0	0	90	95	2	67	1	33
2056	100	100	0	0	0	0	100	100	0	0	0	0

13, the number and percentage of hypercorrection and NL transfer errors is provided for word-initial position and at morpheme junctures before -*ing* and -*y*. These two phonological environments represent the prosodically most salient positions for maintaining the contrast in question.[4] The percentage of hypercorrection errors and NL transfer errors sum to 100. We assumed a threshold of 80% to determine whether a participant's IL grammar should be credited with having the contrast, meaning that the participant's performance had to reach at least the 80% level for each segment in a given environment in order for the IL to be credited with the contrast in that environment.

Thus, for example, Table 1 shows that participant 2038 did not evidence a contrast between [s] and [š] in word-initial position before a high front vowel, because the performance on initial [s] was only 10% of target-like. This participant also did not maintain the contrast intermorphemically because only five percent of the pronunciations of [s] in this environment were target-like. We note in addition that 2038 produced 100% and 90% of the [š] productions as target-like initially and intermorphemically, respectively. Table 1 also indicates that 2038 produced a total of 30 pronunciation errors on [s] and [š], 28 of which were NL transfer errors, and two of which were classified as hypercorrection errors. This resulted in 30% of all errors produced by this participant being NL transfer errors in word-initial position. In the intermorphemic environment, two of the 30 total errors were hypercorrection errors; this amounted to six percent of the total errors being hypercorrections. Finally for 2038, 64% of all errors produced by this participant were NL transfer errors committed in the intermorphemic environment. These results indicate that 2038 was not able to suppress the application of the NL allophonic rule in the production of TL words.

Table 3. Participants' production expressed as percentages on [f] and [v] before a high front vowel word initially and intermorphemically along with numbers and percentages of NL transfer and hypercorrection errors.

Participants	Environments											
	Word-initial (__ i)						Intermorphemic (__ + i)					
	[f] %	[v] %	HC Errors		NL Errors		[f] %	[v] %	HC Errors		NL Errors	
			n	%	n	%			n	%	n	%
2075	10	10	1	13	8	87	0	50	0	0	0	0
2072	30	40	0	0	13	72	95	80	0	0	5	28
2076	50	0	0	0	5	15	30	25	0	0	29	85
2077	20	0	0	0	18	33	0	30	0	0	36	67
2073	60	0	0	0	14	56	75	70	0	0	11	44
2078	80	20	0	0	10	31	45	45	0	0	22	69
2079	80	90	0	0	3	21	95	50	0	0	11	79
2071	70	50	2	10	6	30	85	60	0	0	12	60
2074	80	20	1	5	9	45	90	70	1	5	9	45
2070	100	10	1	6	8	47	95	65	0	0	8	47
2080	100	90	0	0	1	50	100	95	0	0	1	50
2081	100	100	0	0	0	0	100	95	0	0	1	100
2082	100	90	0	0	1	11	90	70	0	0	8	89
2083	70	50	0	0	8	44	80	70	0	0	10	56
2084	100	70	0	0	3	100	100	100	0	0	0	0

Tables 2 and 3 below are structured parallel to Table 1, but represent the participants' performance on the /p/–/f/ and /f/–/v/ contrast, respectively. We elicited the participants' productions of the /p/–/f/ and /f/–/v/ contrasts in word-initial position and in the environment before the suffixes *-ing* and *-y* in order to make the participants' productions on these contrasts parallel to those for /s/ and /š/. In the case of the /s/–/š/ contrast the application of an NL allophonic rule is invoked, whereas with the /p/–/f/ contrast the rule is context-free and in the case of the /f/–/v/ contrast no NL rule is implicated at all.

Considering the data in the tables from the standpoint of the hypotheses in (3) and the facts listed in (4), we note several trends in support of our claims. First, with respect to hypothesis (5a), we can see that it is only the participants in Tables 1 and 2, those producing words with the /s/–/š/ and the /p/–/f/ contrasts, respectively, who systematically evince hypercorrection errors. The participants in Table 1 produced 21% (65 out 315) of the errors as hypercorrections; the participants in Table 2 rendered 24% (25 of 106) hypercorrection errors. We note, however, that 15 out of the 25 hypercorrection errors were made by one participant, 2055, and that the percentage of hypercorrection errors drops to nine percent (10 of 106) if 2055 is eliminated as an outlier. Finally, the participants in Table 3, who produced words containing the /f/–/v/ contrast, do not systematically produce hypercorrection errors, yielding hypercorrections at a rate of only 2.5% (7 of 278). These data support (5a), because systematic hypercorrection errors occur in the

acquisition of the /s/–/š/ and /p/–/f/ contrasts, where at least one of the relevant contrasting segments in the TL also exists in the NL, but systematic hypercorrection errors do not occur in the case of the /f/–/v/ contrast, where neither segment is present in the NL.

We focus now on the two facts about hypercorrection listed in (4), beginning with (4a), which states that hypercorrection errors occur later, rather than earlier, in the acquisition of the contrast. This trend would manifest itself in our data by the existence of a tendency for those participants who produce hypercorrection errors to be to those participants who have mastered the contrast in the prosodically prominent word-initial and intermorphemic, syllable-initial positions. As we see in Table 1, participants who produced only NL transfer errors, with a single exception (2062), lacked the /s/–/š/ contrast word initially, between morphemes, or both, whereas those who produced hypercorrection errors, either exclusively or along with NL transfer errors, reached or approached the criterial threshold in at least one of the environments.

Pursuing this line of reasoning, we propose that an important measure of mastery of the /s/–/š/ contrast would be how successfully the participants can produce [s] before a high front vowel in the pronunciation of TL words. Stated differently, the ability to suppress the IL application of the NL allophonic rule which turns /s/ to [š] before a high front vowel would be a good measure of how well our participants have acquired the contrast in question. Given this assumption, we would therefore expect that the errors made by the participants who have an accurate rendition of [s] before a high front vowel would likely be hypercorrection errors rather than NL transfer errors. We pursued this idea by calculating correlations between participants' production of [s] before high front vowels and the percentage of their hypercorrection errors as a function of the total number of errors that each participant committed. Consistent with our expectations, the results showed a positive correlation between participants' successful production of [s] before a high front vowel and the percentages of hypercorrection errors both word initially, $r(24) = .36$, $p = .07$ (only marginally not significant) (Figure 1a), and intermorphemically, $r(24) = .67$, $p < .001$ (Figure 1b). The correlation is also significant when the data from the two environments are collapsed, $r(24) = .63$, $p < .01$ (Figure 1c). This result suggests that the errors made by the participants who showed a better rendition of [s] before a high front vowel are more likely to be hypercorrection errors. The results also showed a negative correlation between participants' successful production of [s] before a high front vowel and percentages of NL transfer errors, as the greater percentage of hypercorrection errors indicated a lower percentage of NL transfer errors (recall that the percentages of hypercorrection and NL transfer errors sum up to 100%). Overall, these results strongly suggest that the occurrence of hypercorrection and NL transfer errors is not random; rather, they are systematically correlated with the speakers' performance on the TL phonemes and could explain the speakers' TL proficiency.

The results in Figure 2 further show a negative correlation between participants' successful production of [š] before a high front vowel, and percentages of hypercorrection errors both word initially, $r(24) = -.53$, $p < .01$ (Figure 2a), and intermorphemically, $r(24) = -.62$, $p < .01$ (Figure 2b). The correlation was also significant when the data from the two environments were combined, $r(24) = -.60$, $p < .01$ (Figure 2c).

This same general pattern holds also for Table 2, describing the participants' performance on the /p/–/f/ contrast. Those participants (2052, 2024) that made only NL transfer
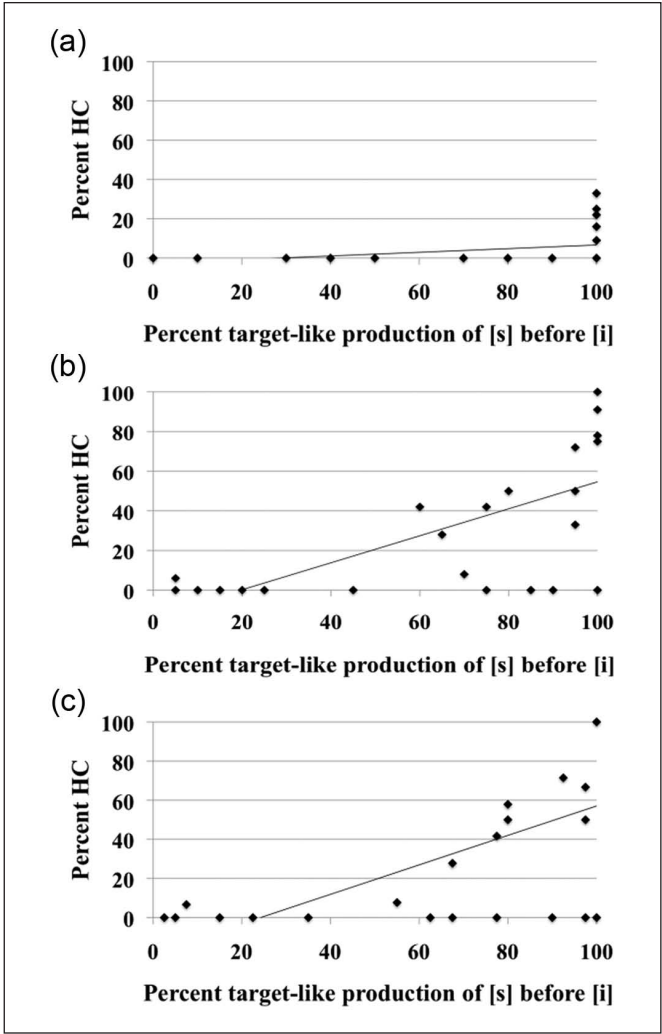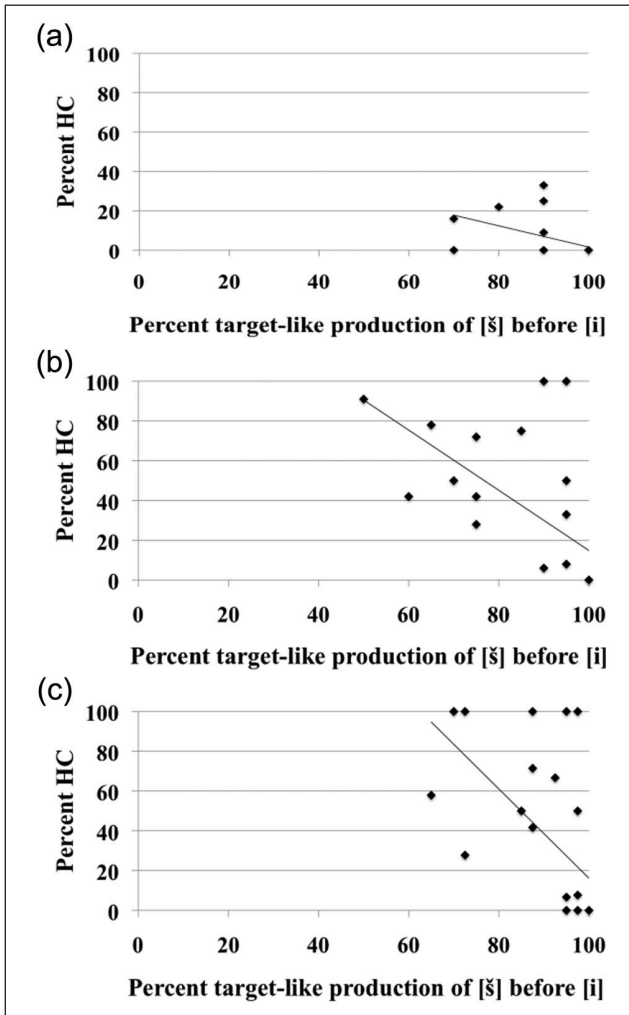
Figure 1. The relationship between percent target-like production of [s] before [i] and percent hypercorrection (HC), (i.e. substitution of [s] for TL [š] before [i]) word-initially (a), inter-morphemically (b), and in both environments (c)).

errors both lack the contrast on /p/ and /f/ between morphemes. The participants who produced both types of errors (2044, 2042, 2055) showed the contrast in at least one of the prominent positions. This supports the claim that hypercorrection errors occur later, rather than earlier, in the acquisition process, and at the very least and as a matter of necessity, hypercorrection errors can occur only after a learner has shown evidence of the contrast, at least to some extent.

Parallel to the contrast between /s/ and /š/, the successful rendition of [f], a non-native phoneme in Korean, would be an indication of a Korean learner's mastery of the
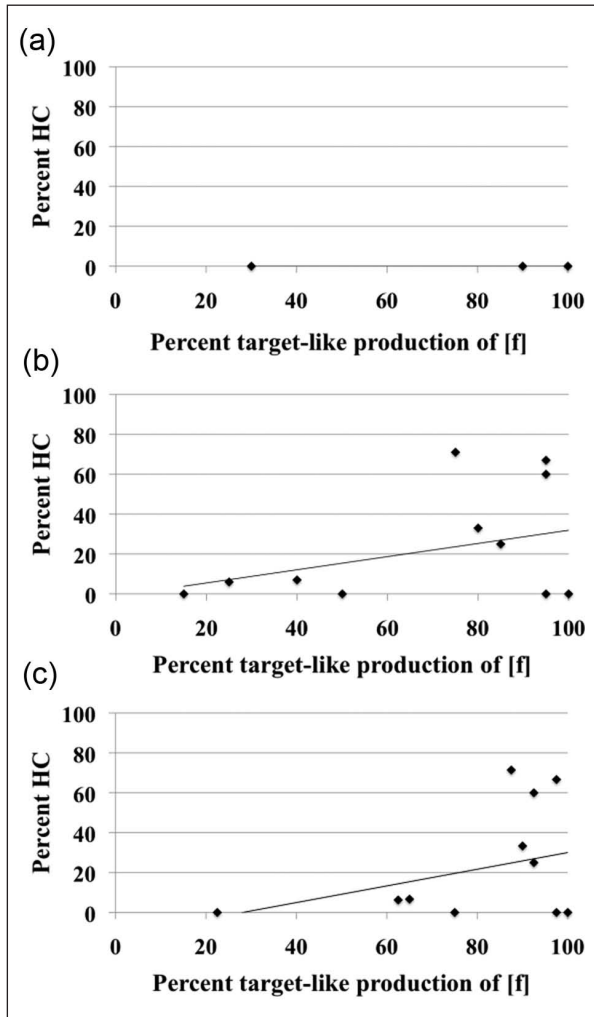
Figure 2. The relationship between percent target-like production of [š] and percent hypercorrection (HC), (i.e. substitution of [s] for TL [š] before [i]) word-initially (a), inter-morphemically (b), and in both environments (c)).

English /p/–/f/ contrast. Thus, we would expect to find a positive correlation between Korean speakers' target-like production of [f] and the occurrence of hypercorrection errors. That is, errors made by the participants who produce [f] successfully are more likely to be hypercorrection errors than NL transfer errors. We were not able to investigate a correlation between the two in word-initial position, as there were no hypercorrection errors in this environment (Figure 3a). In intermorphemic position, we did not find a correlation between Korean speakers' production of [f] and the occurrence of hypercorrection errors, $r(10) = .37$, $p = .256$ (Figure 3b). Moreover, this correlation was

Figure 3. The relationship between percent target-like production of [f] and percent hypercorrection (HC), (i.e. substitution of [f] for TL [p]) word-initially (a), inter-morphemically (b), and in both environments (c)).

also not significant when the data from the two environments were combined, $r(10) = .33$, $p = .288$ (Figure 3c). We should also acknowledge that these results include partici- pant 2055, who is responsible for fully 60% of the hypercorrection errors. If we remove his results from the calculations, the correlation in the combined environments is still not statistically significant, $r(9) = .35$, $p = .293$. Nevertheless, note that the correlation between participants' performance on [f] and hypercorrection errors is in the right direction (i.e. positive relationship), and perhaps with more statistical power gained from a greater number of participants, the correlation would be statistically reliable.
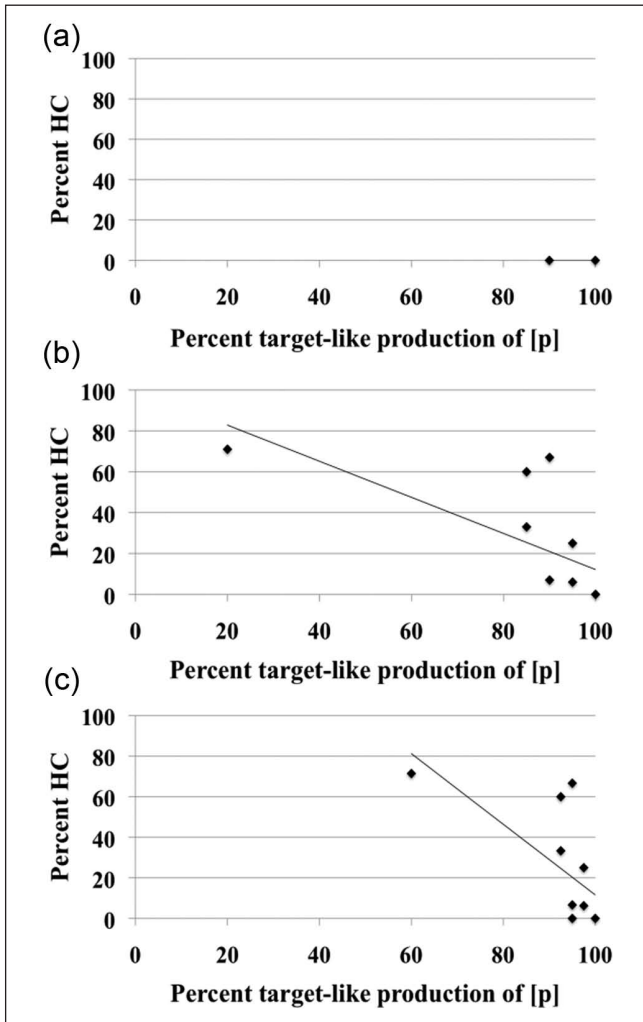
Figure 4. The relationship between percent target-like production of [p] and percent hypercorrection (HC), (i.e. substitution of [f] for TL [p]) word-initially (a), inter-morphemically (b), and in both environments (c)).

As mentioned above, we were not able to investigate a correlation between participants' successful production of [p] and percentages of hypercorrection errors as a function of the total number of errors in word-initial position, because there were no hypercorrection errors in word-initial position (Figure 4a). However, the results did show a significant negative correlation between these two measures in intermorphemic position, $r(10) = -.69$, $p < .05$ (Figure 4b), though one of the participants (2055) seems to be an outlier, producing 15 of the total 35 hypercorrection errors for the participants producing the /p/–/f/ contrast. The correlation was also significant when the data from the two environments were combined, $r(10) = -.67$, $p < .05$ (Figure 4c).
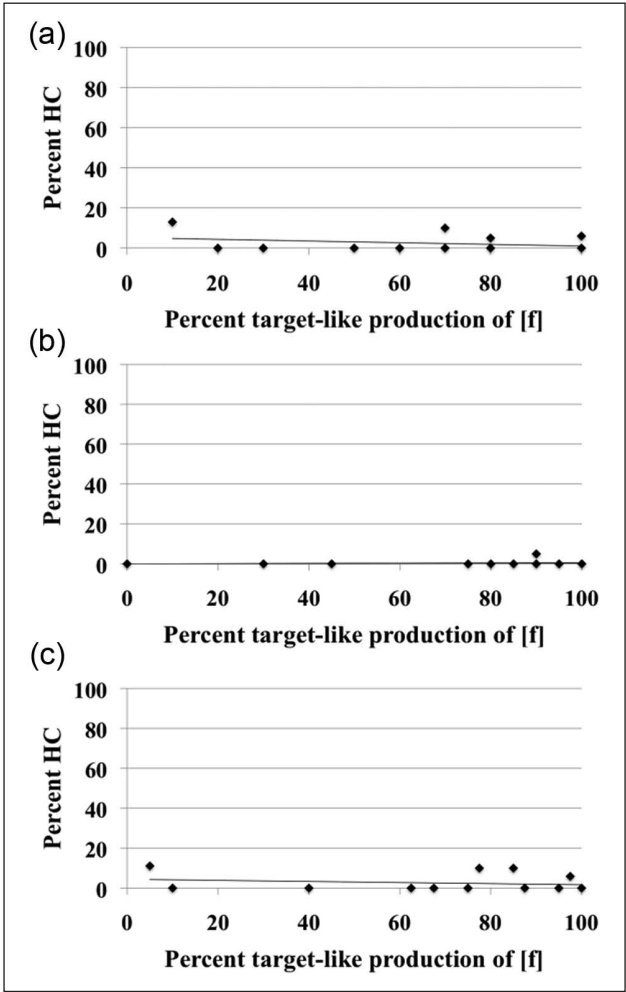
Figure 5. The relationship between percent target-like production of [f] and percent hypercorrection (HC) word-initially (a), inter-morphemically (b), and in both environments (c).

Finally, we turn to the results of our participants' production of the /f/–/v/ contrast. As predicted by the hypotheses in (3), a correlation between target-like production of [f] and percentages of hypercorrection errors was not significant in either environment: word initially, $r(13) = -.30$, $p = .272$ (Figure 5a), and intermorphemically, $r(13) = .14$, $p = .619$ (Figure 5b). Moreover, this correlation was also not significant when the data from the two environments were combined, $r(13) = -.20$, $p = .484$ (Figure 5c)

We also did not find a relationship between the target-like production of [v] and percentages of hypercorrection errors: word initially, $r(13) = -.30$, $p = .281$ (Figure 6a), or intermorphemically, $r(13) = .06$, $p = .827$ (Figure 6b). Again, this correlation was not significant when the data from the two environments were combined, $r(9) = -.26$, $p = .343$ (Figure 6c).
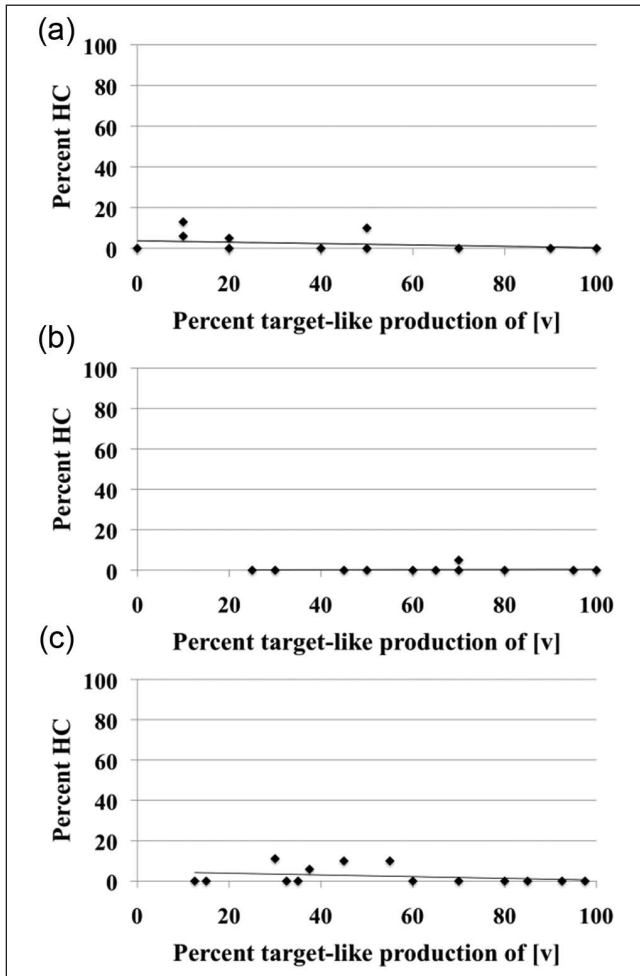
Figure 6. The relationship between percent target-like production of [v] and percent hypercorrection (HC) word-initially (a), inter-morphemically (b), and in both environments (c).

Thus, the data in Tables 1 through 3 and the correlations in Figures 1 through 6 support the claim that NL transfer errors and hypercorrection errors are not random, but appear to be related to a participant's performance on the relevant contrast, with the percentage of hypercorrection errors being the greatest where an NL context-sensitive rule is involved, less when the NL has only one of the segments in the TL contrast being learned, and virtually absent when the NL contains neither of the TL sounds being contrasted.

To summarize this section, our results support the hypotheses in (5), which claim that at least part of the mechanism for hypercorrection errors is the transfer of some NL rule or constraint into the IL, where the unlearning or suppression of that rule or constraint forms the basis for hypercorrection errors.

V   Discussion

In this article we have argued that hypercorrection in SLA cannot be straightforwardly characterized as crosslinguistic overgeneralization errors because, on the one hand, they are not over-extensions of an NL pattern, and, on the other hand, some NL–TL differences do not lead to hypercorrection errors. Within this context, we wish to discuss three points pertaining to the relevance of hypercorrection errors to our understanding SLA.

Our first point harks back to the language contact situation described in (1c) above, namely the acquisition of a TL phonemic contrast in which the NL has both of the sounds in question, but as allophones of the same phoneme. The relative difficulty involved in acquiring such an allophonic split was first discussed, to the best of our knowledge, by Lado (1957: 15), who described this learning task as representing maximum difficulty for L2 learners. Given Lado's framework, in which NL–TL differences are paramount for explaining learning difficulty, it is surprising that an allophonic split (where the NL and TL are similar with respect to having the sounds in question) would constitute more difficulty than, say, the contact situation in (1a), where the NL has neither segment in the TL contrast. Although Lado himself provided only anecdotal evidence for this conclusion, studies over the decades have borne out his claim. Thus, Stockwell and Bowen (1965) rank allophonic splits high in their hierarchy of phonological difficulty, as do Hammerly (1982) and Hardy (1993) in their empirical studies. The relative difficulty of this kind of learning has also been attested for misarticulating children by Gierut (1986, 1988).

We suggest that our findings may shed some light on understanding the reported intractability of this learning situation. As analysis of the data in Table 1 shows, hypercorrection errors tend to be produced by participants who have mastered, at least to some extent, the contrast in question. Moreover, the correlations in Figure 1 indicate that the relationship between performance on the contrast and the production of hypercorrection errors is positively correlated and statistically reliable in the case of the phonemic split for /s/ and /š/. In fact, it would not be unreasonable to conclude on the basis of this evidence that the production of hypercorrections may well be the final stage of acquiring a TL phonemic contrast. This claim is supported by two facts. The first is that a learner must necessarily have the contrast in question, at least to some extent, before the contrast can be hypercorrected; indeed, this would seem to be a logical necessity. The second is that the results in Table 1 attest a group of participants who produce hypercorrection errors, but no NL transfer errors. Moreover, given the prevailing view that hypercorrection errors stem from a certain amount of linguistic insecurity, on the one hand, and relative prestige being attached to the TL, it would be reasonable to conclude that this error type could potentially emerge with the acquisition of each new lexical item that contained the relevant contrast. As we argued above, we are not able, nor is it the intent of our analysis, to forecast whether or when L2 learners will err. Rather, the claim is that, if learners who already have the contrast in question are going to make errors, the errors are more likely than not to be hypercorrection errors.

It is also clear from our analysis that the acquisition of some types of contrast will not produce hypercorrections. If the NL lacks sounds corresponding to both of the TL phonemes in question, no systematic hypercorrection errors will occur. This was exemplified in the Korean data relating to the productions of the English /f/–/v/ contrast. Our findings

show no hypercorrection in this case because the NL contains neither of the phones in question; thus, since neither [f] nor [v] exists in the NL, there appears to be no basis for a Korean learner to produce one of the segments as a substitution for the other. Indeed, the overwhelming majority of errors produced for English /f/ and /v/ were [p] and [b], respectively.

To conclude this point, we can only speculate on what the evidence was that motivated Lado (1957: 15) to assert that making an allophonic split constituted maximum learning difficulty. However, the facts are clear from our data that not only is the learning of this type of contrast staged (Eckman and Iverson, 2013), but also that errors in mastering this type of contrast are candidates to persist in the form of hypercorrections.

The second point that we wish to discuss is the distinction between 'correction' in the learning of a new TL contrast, and 'hypercorrection'. When an L2 learner is faced with acquiring a new TL distinction, 'correction' in the IL phonology results from the learner producing one or both members of the contrast in environments where the segments do not occur in the NL; 'hypercorrection', on the other hand, results from the learner inappropriately extending those newly learned sounds, or newly learned environments for sounds. As we have touched on, this over-extension results from the linguistic insecurity involved with learning new contrasts, and the prestige associated with the TL. Thus, correction in learning the /p/–/f/ contrast by Korean learners of English comes from learners producing the new TL sound, [f], in all relevant positions; hypercorrection stems from extending the production of this new sounds to positions where it should not occur. In the same vein, correct learning of the English /s/–/š/ contrast by native speakers of Korean requires implementing [s] before high front vowels and vocoids, whereas hypercorrecting this contrast stems from incorrectly extending the pronunciation of [s] in this environment when it should not occur.

The final topic we wish to discuss is related to the point made in the preceding paragraph, and causes us to recapitulate hypothesis (5b), namely that when a context-sensitive rule is transferred into the IL, hypercorrection errors occur predominantly, if not exclusively, in the environment of that rule. Thus, for example, all of the hypercorrection errors listed in Table 1 occurred before a high front vowel, the environment specified in the transferred Korean allophonic rule. The potential for a hypercorrection error in an environment other than before a high front vowel arises in our data in intervocalic position in mono-morphemic words. The target words in question are listed in (6) below.

(6)   Target words containing an intervocalic [š] (American English)

ocean

parachute

patient

vacation

tissue

As only the word *tissue* in (6) contains a high front vocoid, and since none of the words contains [š] before a high front vowel, they provide an opportunity for participants to produce a hypercorrection in an environment other than that specified in the rule in

(2). A total of seven participants – 2027, 2036, 2037, 2038, 2047, 2049 and 2051 – produced a total of 11 errors on medial [š] on the words in (6). Of those 11 errors, only two can be classified as a possible hypercorrection error, participant 2027's two productions of *parachute* with a medial [s]; in all other cases, the error involved the production of a segment other than [s]. Thus, our claims about the phonological context in which hypercorrection errors will occur is further supported.

VI   Conclusions

This article has reported the results of a study on the NL transfer and hypercorrection error types found in the acquisition of three kinds of L2 phonemic contrasts: the splitting of NL allophones into phonemes, the learning of a TL segment not contained in the NL, and the acquiring of a contrast between two segments neither of which is found in the NL. Our findings suggest that the production of hypercorrection errors may be the final stage in the acquisition of a TL phonemic contrast, and that the occurrence of such errors is based on the strength of the connection between the TL contrast being acquired and the learner's NL phonology.

Notes

1. The results on the acquisition of the /s/–/š/ contrast for these participants has been reported as part of a separate study by Eckman and Iverson, (2013). The participants were all students in the University of Wisconsin-Milwaukee Intensive English Program. Although the vast majority of the participants in this project could be reasonably termed 'intermediate' or 'intermediate high' in their ability in English, we did not conduct any measure to determine the English proficiency of the participants.
2. The results on the acquisition of the production and perception of the /p/–/f/ contrast for a subset of these participants has been reported elsewhere as part of a different study by (Eckman and Iverson, 2011).
3. We chose the words to be elicited on the basis of the relevant phonological environment and the relative ease with which the word could be elicited from an image displayed on a computer screen. Paramount to us was that the participants were not exposed to the spelling of the word, as this could have materially affected the participants' pronunciation. An anonymous reviewer pointed out that, on the basis of studies by Bybee (2000, 2001), the frequency of the elicited words could have affected the likelihood that the participants produced hypercorrection errors. As we have no data bearing on this matter, we must leave the question open.
4. For considerations of clarity and ease of exposition, we do not present the participants' scores on the contrasts in word-medial or word-final position, as these results are less indicative

of participants' performance on the contrasts. For the /s/–/š/ contrast, only eight participants performed below 100% in either medial or final position, and only one of those participants scored lower than 80%. Moreover, the words that contained a medial [s] or [š] did not have these segments before high front vowels, as it was impossible to find readily picturable English words with [s] or [š] in this environment. The usefulness of these words with medial [s] and [š] comes in determining whether any hypercorrection errors occurred in environments other than before a high front vowel.

References

Bybee J (2000) The phonology of the lexicon: Evidence from lexical diffusion. In: Barlow M and Kemmer S (eds) *Usage-based models of language*. Stanford, CA: Center for the Study of Language and Information, 65–85.

Bybee J (2001) *Phonology and language use*. Cambridge: Cambridge University Press.

Campbell L (1998) *Historical linguistics: An introduction*. Cambridge: The MIT Press.

Eckman FR, Iverson GK, Fox RA, Jacewicz E and Lee SA (2011) Explicit training and implicit learning of L2 phonemic contrasts. In: Sanz C and Leow RP (eds) *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*. Washington, DC: Georgetown University Press, 159–74.

Eckman F and Iverson GK (2013) The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition* 35: 1–26.

Gierut JA (1986) Sound change: A phonemic split in a misarticulating child. *Applied Psycholinguistics* 7: 57–68.

Gierut JA (1988) Comparative research on language learning. *Language Learning* 38: 413–38.

Hammerly H (1982) Contrastive phonology and error analysis. *International Review of Applied Linguistics* 20: 17–32.

Hardy J (1993) Phonological learning and retention in second language acquisition. In: Eckman F (ed.) *Confluence: Linguistics, L2 acquisition and speech pathology*. Amsterdam/Philadelphia: John Benjamins, 235–48.

Hecht BF and Mulford R (1982) The acquisition of a second language phonology: Interaction of transfer and developmental factors. *Applied Psycholinguistics* 3: 313–28.

Janda R and Auger J (1992) Quantitative evidence, qualitative hypercorrection, sociolinguistic variables – and French speakers' *'eadhaches* with English h/Ø. *Language and Communication* 12: 195–236.

Johansson FA (1973). Immigrant Swedish phonology. Lund: Gleerup.

John P and Cardoso W (2009) Francophone ESL learners' difficulties with English /h/. In: Baptista B, Rauber A and Watkins M (eds) *Recent research in second language phonetics/phonology: Perception and production*. Newcastle Upon Tyne: Cambridge Scholars, 118–40.

Kim H (2001) The place of articulation of the Korean plain affricate in intervocalic position: An articulatory and acoustic study. *Journal of the International Phonetic Association* 31: 229–57.

Knowles G (1978) The nature of phonological variables in Scouse. In: Trudgill P (ed.) *Sociolinguistic patterns in British English*. London: Edward Arnold, 80–90.

Kohler K (1971) On the in adequacy of phonological theories for contrastive studies. In: Nickel G (ed.) *Papers in contrastive linguistics*. New York: Cambridge University Press, 83–88.

Labov W (1963) The social motivation of a sound change. *Word* 19: 273–309.

Labov W (1966) *The social stratification of English in New York City*. Washington, DC: The Center for Applied Linguistics.

Labov W (1972) *Sociolinguistic patterns*. Philadelphia, PA: The University of Pennsylvania Press.

Lado R (1957) *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

Major R (2001) *Foreign accent*. Mahwah, NJ: Lawrence Erlbaum.

Nemser W (1971) *An experimental study of phonological interference in the English of Hungarians*. Bloomington, IN: Indiana University Press.

Odlin T (1989) *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Siegel J (2003) Social context. In: Doughty CJ and Long MD (eds) *The handbook of second language acquisition*. Malden, MA: Blackwell, 178–223.

Stockwell R and Bowen D (1965) *The sounds of English and Spanish*. Chicago, IL: The University of Chicago Press.

Wolfram W (1991) *Dialects and American English*. Englewood Cliffs, NJ: Prentice Hall.

Appendix 1. Target words used for eliciting the [s]–[š] contrast.

1.	scissors	31.	wish
2.	sick	32.	bush
3.	seats	33.	wash
4.	sink	34.	fish
5.	see	35.	polish
6.	kiss	36.	splash
7.	race	37.	rush
8.	mess	38.	crush
9.	pass	39.	brush
10.	price	40.	crash
11.	dress	41.	racing
12.	grass	42.	kissing
13.	cross	43.	messing
14.	bus	44.	passing
15.	face	45.	pricing
16.	lesson	46.	pricey
17.	message	47.	dressy
18.	bicycle	48.	grassy
19.	motorcycle	49.	crossing
20.	license	50.	bussing
21.	she	51.	bushy
22.	sheet	52.	washing
23.	ship	53.	splashing
24.	sheep	54.	polishing
25.	shield	55.	rushing
26.	ocean	56.	wishing
27.	tissue	57.	crushing
28.	vacation	58.	fishing
29.	patient	59.	brushing
30.	parachute	60.	crashing

Appendix 2. Target words used for eliciting the [p]–[f] contrast.

1. pin	31. leaf
2. pot	32. stuff
3. panda	33. buff
4. pen	34. cough
5. pie	35. knife
6. foot	36. beef
7. face	37. laugh
8. fire	38. puff
9. fish	39. graph
10. farm	40. stuff
11. paper	41. sleepy
12. zipper	42. sleeping
13. report	43. shopping
14. happy	44. soapy
15. Nepal	45. hopping
16. before	46. mopping
17. coffee	47. dropping
18. telephone	48. shipping
19. professor	49. typing
20. cafe	50. rapping
21. leap	51. buffing
22. sleep	52. knifing
23. drop	53. coughing
24. soap	54. puffing
25 . ship	55. stuffing
26 . shop	56. leafy
27. mop	57. beefy
28. hop	58. graphing
29. type	59. laughing
30. rap	60. handcuffing

Appendix 3. Target words used for eliciting the [f]–[v] contrast.

1. face	31. give
2. fish	32. save
3. fan	33. move
4. foot	34. love
5. fire	35. arrive
6. violin	36. beefy
7. vine	37. puffing
8. vase	38. graphing
9. van	39. stuffing
10. vote	40. coughing
11. oven	41. handcuffing
12. heaven	42. buffing
13. seven	43. knifing
14. river	44. laughing
15. devil	45. leafy
16. handcuff	46. waving
17. leaf	47. diving
18. beef	48. shaving
19. puff	49. giving
20. graph	50. moving
21. stuff	51. coffee
22. cough	52. telephone
23. laugh	53. cafe
24. buff	54. buffalo
25. knife	55. referee
26. wave	56. living
27. dive	57. driving
28. shave	58. saving
29. live	59. loving
30. drive	60. arriving