II. 1. High Vowel Devoicing (HVD)

HVD devoices high vowels, i and u, when they occur between voiceless consonants or before a pause. Although the formalization of this process is not so simple in slow or careful speech, it is generalized and simplified in a fast speech context, which is something like (1).

(1) High Vowel Devoicing (HVD)

\[
\begin{align*}
\{\text{[y1] high}\} & \rightarrow \{\text{[y1] voiceless}\} & \{\text{[y1] voiceless}\} \rightarrow \{\text{[y1] pause}\}
\end{align*}
\]

This rule generates the following phonetic representations, where Y stands for a devoiced vowel and V for a long vowel.

(a) kṣety 'season'
(b) kṣetu ga 'season SM'
(c) sykos 'a little'
(d) sykos da 'only a little'
(e) yakusoku 'promise'
(f) yakusoku da 'promise copula'
(g) tij[k]a[k]yū 'intelligentsia'
(h) tōkyō e 'kyōhi ro
(i) tōkyō e iku basy

As is shown above, HVD applies regardless of word boundaries or syntactic configurations. It is purely conditioned by phonological environment and a certain rate of speech production.

II. 2. Vowel Degemination and Lengthening (VDAL)

Japanese has five short vowels, i, e, a, o, u, and corresponding long vowels, I, E, A, O, U, in both underlying and surface levels. A surface long vowel can be created by two identical short vowels. Thus, there is not any phonetic distinction between the pairs, such as sūri 'mathematical principle' and sūrīri 'vined brewer', satōya 'sugar salesman' and sato'ya 'foster parent'. This process can be written as (2), which degeminates two identical vowels and changes them into a longer vowel.

(2) Vowel Degemination and Lengthening (VDAL)

\[
\begin{align*}
\text{V_1} \rightarrow \text{V_1} \rightarrow \text{V_1}
\end{align*}
\]

This process is not confined to a word level but it applies anywhere given a certain rate of speech. Observe below:

(a) ka+i+n → ka+i+n 'member'
(b) akurun → akun 'devil's luck'
(c) aki le → aki e 'red house'
(d) hako → hako 'box OM'
(e) uti isou → utiisou 'should too early'
(f) tōkyō ni iku → tōkyō ni iku 'go to Tokyo'
(g) goku ga atta → goku gatta
(h) tari ni i o yaru → tari ni i o yaru 'I met'
(i) tari ni i o yaru → tari ni o yaru
(j) satōya → satōya 'sugar salesman'
(k) sato'ya → sato'ya 'foster parent'

In this section, I will take up two rules which are sensitive to the rate of speech. They are incidentally neutral to sociological and lexical factors.
The above two processes share the following common features: (i) the occurrence of these processes are totally dependent on the rate of speech; (ii) their operation is not sensitive to lexical items, the casualness of the situation where the speech takes place, or the syntactic configurations; (iii) they are phonetically natural—NVD is a voicing assimilation rule which affects less sonorous vowels and VD&L is a lenition process by which the effort to utter two identical sounds separately is eased.

III. Casual Speech Processes

Three different types of casual speech processes will be presented in this section, all of which can take place in a slow or careful speech as well as in a fast speech. Where they occur seems to be sensitive to the notion of 'casualness'; how casual or informal the situation is has to be specified for a certain process to be triggered.

III. 1. Nasal Syllabification (NS)

When no, either as a nominalizer (5a) or as a possessive marker (5b), occurs before a copula, g in no drops and n is syllabified. Note that the copula always starts with either n, as in ni 'connective', nara 'provisional', or d, as in da, desu 'present', datta, deshita 'past', daro, deso 'tentative'. Similarly, the connective form of the copula, ni, becomes g, being followed by nar 'become'. Such processes are apparently related to the identity of the place of articulation of n or d of ni and the following consonant, d or n, of the copula and nar. It has to be noted that NS never occurs if n or d following after the no or ni in question is not part of a copula or nar.

Observe the following:

(5a) a. kuru no nara → kuru g nara come NOM. Copula-cond. 'If it is that... come...'  
   b. taro no da → taro g da Taro Poss. Copula-pres. 'It is Taro's.'  
   c. kimi no namae → kimi g namae you Poss. name 'your name'  
   d. gaku na ni naru → gaku ga na naru scholar Copula-conne. become 'become a scholar'  
   e. genki ni naru → genki g naru fine Copula-conne. become 'become fine'  

The sociological conditioning of this rule is not terribly strong. The syllabized form g can occur in a fairly formal or polite expression, regardless of the rate of speech. However, the opposite is not true. The occurrence of the full form does not seem to be appropriate in the vulgar expression. For instance:

(6a) tenne ga yatta (?) da.
   you-vulgar SM did NOM. Copula 'It was that you did it.'

In (6), the occurrence of the lexical item tenne 'you', which is a very vulgar form of the second person pronoun, forces the entire string to be marked as 'vulgar'. In such an expression, NS seems to be obligatory, leaving the casual form g.

The following is the tentative formalization of this rule, which is equipped with all the lexical information discussed above.

(7) Nasal Syllabification (NS)

\[
\{\text{n} \} \rightarrow \text{g} / \quad \{\text{copula}\}
\]

Cond.: ni = copula

III. 2. Vowel Fusion (VF)

As compared to NS, Vowel Fusion (VF) takes place under much more restricted sociological conditions. In men's vulgar speech only, ai, at, and ae are fused into a long vowel å. The process can be formalized as (8).

(8) Vowel Fusion (VF)

\[
\{\text{ai}\} \rightarrow \text{å} / \quad \{\text{ae}\}
\]

This rule does not apply across words or phrases (see (9f) and (9g)).

(9) a. atarimae → atarime 'of course'
   b. kāru → kāru 'return'
   c. sugoi → sugē 'great'
   d. Verb-tai → Verb-TE 'want to Verb'
   e. nai → nē 'not, no'
   f. ore gaiku → ore gēku 'I go'
   g. ore noiku → ore mēku 'I also go'

The application of VF seems to be restricted to certain lexical items. For example, the following items, among others, do not seem to undergo this process.

(10) a. kōfu → kōfu 'this man'
   b. kīro → kīro 'yellow'
   c. saigo → saigo 'the last'
   d. kimae → kimē 'generosity'

Furthermore, it has to be noted that VF can be regarded as a dialectal process, to the extent that it is much more frequently observed in the speech of a certain social group, e.g., gangsters, the scum of the streets among the Tokyo dialect speakers, and of the speakers of certain regional dialects. Nonetheless, fused forms of tai 'want to', nai 'not, no', fixed expressions, such as (9a), and the ending of commonly used adjectives, such as sugō 'great', semai 'narrow; small' are fairly common in men's speech, if the situation allows the occurrence of unrefined speech.

Although the phonological nature of VF is to reduce the energy for producing two different vowels, its application would never be over extended by merely accelerating the rate of speech. Thus, one may speculate that the rigidity of the sociological condition does
III. 3. Two Contraction Processes

Among several contraction processes exhibited in Japanese, I will take up the following two, which involve the syntactically rather complex strings. Both of them have to be strictly conditioned by lexical information. See the following examples:

(11) a. mi-te $imagin-u$ $\rightarrow$ mi-$imagin$-$u$ $\rightarrow$ mi-$caw$-$u$
    see-Gerun. finish-pres. 'have seen'
    b. yom-de $imagin-u$ $\rightarrow$ yom-$imagin$-$u$ $\rightarrow$ yom-$jaw$-$u$
    read-Gerun. finish-pres. 'have read'
    c. kore-de $imagin$-$imasu$ $\rightarrow$ kore-$imagin$-$imasu$ $\rightarrow$ kore-$jaw$-$imasu$
    this at finish-polite-pres. 'finish at this moment'
    d. mi-te $shi$-$ru$ $\rightarrow$ mi-$ci$-$ru$ $\rightarrow$ mi-$cu$
    see-Gerun. know-pres. 'know by seeing'
    e. yom-de $shi$-$ru$ $\rightarrow$ yom-$fj$-$ru$ $\rightarrow$ yom-$fu$
    read-Gerun. know-pres. 'know by reading'

(12) a. mi-$ru$ $\rightarrow$ koto wa (nai) $\rightarrow$ mi-kko (nai)
    see-pres. NOM. not not see-infin.
    'It is not the case that ... see ...
    b. yom-$ru$ $\rightarrow$ koto wa (nai) $\rightarrow$ yom-i kko (nai)
    read-přes. NOM. not not read-infin.
    'It is not the case that ... read ...
    c. mi-$ta$ $\rightarrow$ koto wa (nai) $\rightarrow$ mi-ta kko (nai)
    see-past $\rightarrow$ mi-kko (nai)
    'have not seen ...'

In (11), it is shown that the contractable string must involve the gerundive marker te (de) and the aspectual verb $imagin$. Note that de as an adverb ending and $imagin$ as an independent verb cannot be contracted (see (11)c), nor can a phonetically similar string, such as (11d) and (11e). Likewise, only Verb-Pres. $\rightarrow$ koto wa (nai) 'it is not the case that ... Verb ...' changes into Verb-infinitive kko (nai), as shown in (12). Thus, these two processes involve only these lexically and syntactically fixed items. I tentatively give the following rules.

(13) a. (Verb)-to (de) $imagin$ $\rightarrow$ $imagin$ (Jimaw) $\rightarrow$ $jaw$ (Jaw)
    Gerun. finish
    b. Verb-Stem (r)u $\rightarrow$ koto wa (nai) $\rightarrow$ Verb-Infin. kko (nai)
    pres. $\rightarrow$ Nom. not

The above contracted forms do occur in fairly formal speech, although the situation has to be marked as 'intimate'. They tend to give a rather childish color to the entire conversation. In fact, the contracted forms are extremely popular in children's speech. In addition, these two processes are not at all sensitive to the rate of speech; no matter how fast the speech is produced, the full form would never be contracted. I will return to these processes in the next section.

IV. Casual Speech Processes and the Lexicon.

As I first mentioned, the distinction between FS and CS processes is primarily set depending on whether the rate of speech is relevant to a process. CS processes are insensitive to the speed of utterance production. Besides this, there are common characteristics in the CS rules described above, which are not shared by the FS rules: (1) CS rules are not or less conditioned by lexical information; (2) CS rules are sensitive to sociological notions. These two are not only the traits of CS phonological processes but those of CS morphological and syntactic processes. This means that lexical conditioning is a general characteristic of processes sensitive to the sociological notions. Let me illustrate this point by taking up the following rules as representative: 0 or Go Attachment (0/Go/A)—a morphological process, Subject Honorification (SH)—a syntactic process.

A/Go/A is a morphological process which prefixes o or go to a noun or an infinitive verb in a polite speech (especially women's). It is often said that o is prefixed to a native word and go to a Sino-Japanese word. To the extent that this process has a number of exceptions, it has to be conditioned lexically; some words would not undergo this process (see (14a, b, and c)), some Sino-Japanese words do not choose go but o (see (14d, f, g, and h)).

(14) a. sekken $\rightarrow$ go-sekken 'soap, S-J'
    b. tenugui $\rightarrow$ go-tenugui 'towel, N'
    c. kudarono $\rightarrow$ go-kudarono 'fruit, N'
    d. yasai $\rightarrow$ go-yasai o-yasai 'vegetable, S-J'
    e. shitorei $\rightarrow$ go-shitorei o-shitorei 'impolette, S-J'
    f. burrei $\rightarrow$ go-burei o-burei 'impolette, S-J'
    g. safin $\rightarrow$ go-safin o-safin 'photo, S-J'
    h. sokujii $\rightarrow$ go-sokujii o-sokujii 'meal, S-J'

Besides these exceptions, it has to be noted that some are more resistant than others to undergoing this process and that some prefixed forms become semantically independent items (e.g., go-han 'meal, rice' or-kaze 'thanks to, owing to', etc.).

The syntactic process, SH, which in general attaches o or go in front of a verb (the same operation as O/Go/A), changes the verb into the infinitive form, and inserts ni-nar 'copula—become—Honorific verb ending' after the verb, is also lexically governed to the extent that certain lexical items do not appear in the string, o (or go) Verb-Infin. ni-nar, which is created by SH. Such items possess a suppletive form, which replaces the string made by SH.

(15) a. kak- 'write' o-kakki-ni-nar 
    b. iw- 'say' o-iti-ni-nar but ossar
    c. ik- 'go' o-iti-ni-nar but o-ide-ni-nar
    d. mi- 'see' o-mi-ni-nar but go-rag-ni-nar

Just like the above morphological and syntactic processes of CS phenomena, the phonological CS rules are also lexically conditioned as discussed in the former section. This does not seem to be a mere
coincidence but can be regarded as a significant characteristic of CS phenomena in general. In fact, the lexicon is the most sensitive to such sociological notions. For example, the first person singular pronoun can be any of the following (and perhaps more) depending on the sociological context:

(16) a. wataku'shi 'formal'
   e. boku 'general-boys, young men'
   b. wata'shi 'general'
   f. ore 'informal-men'
   c. ataku'shi 'formal-women'
   g. wa'shi 'informal-older men'
   d. atashi 'general-women'
   h. a'shis with certain occu-

The generalization here is that CS processes are lexically governed, regardless of components of grammar, lexical, morphological, syntactic, or phonological. This has been completely ignored in the generally accepted model of the grammar of Japanese, where each time a process sensitive to sociological notions is operative, lexical information has to be referred to. Therefore, it is a mystery why lexical information has to be supplied whenever a process sensitive to sociological condition is operating—thus, the missing of a generalization.

On the other hand, if a grammar is so structured that all the CS processes are in the lexicon, such loss of generalization would not result. Furthermore, such a model can restrict the entire structure of a grammar in such a way that sociological conditions are allowed only in the lexicon and not in any other components of a grammar. Hence, this approach to CS phenomena is desirable. In such a model, all the CS phenomena discussed above should be described in the lexicon.

In the case of NS, both a casual (syllabified) form, g, and the full forms, no and ni, are in the lexicon and a copula and nar 'become' are marked to choose either the full form or g. The relationship between the full forms and g can be captured if the lexicon is viewed along the line of Hust (1978), where related lexical items are given in a branching diagram.

As for VF, fused lexical items are given in the lexicon, as well as nonfused ones. This approach is compatible with the following fact: one often means a different thing from a non-fused counterpart. For example, the humble form of the first person pronoun temae becomes the vulgar form of the second person pronoun teme, if it is fused. Thus, it is impossible to derive teme from temae by the phonetic rules of VF, applying after lexical insertion. Furthermore, the existence of teme seems to suppress fusing ome '2nd person pronoun-

informal' into ome '2nd person pronoun-informal' into ome 'in vulgar speech, although the fused form often does exist as a 2nd person pronoun which gives a local or dialectal color to the entire speech. Thus, in the Tokyo dialect, the vulgar counterpart of ome is teme, rather than ome. Such subtle differences in the usage of these forms are best described in the lexicon.

As for 0/G0-A and SH, it seems to be intuitively natural to describe these phenomena in the lexicon. 0/G0-A has to refer to almost every lexical item in order to decide whether o or go (or nothing) is to be attached. Why not then list each item with or without a certain prefix in the lexicon structured along the line of Hust? SH not only involves 0/G0-A but also the insertion of the lexical items, ni and nar, the change of a verb form from the stem to the infinitive, and the replacement of lexical items with their suppletive forms. Although how SH phenomena are captured in the lexicon has to await future researches, the existence of SH as a transformation can be criticized on independent grounds; a grammar which allows the insertion of the full form and/or the replacement of lexical items by transformations is too powerful. Moreover, if SH phenomena can be described in the lexicon, the suppletive form is no longer exceptional but follows naturally from the lexical specification.

The two contraction processes discussed in III. 3. are to be captured in the lexicon. Notice that in (13a) the casual form kko, which is a contracted form of a syntactically complex string Verb-Pres-ci kko wa requires the infinitive form of the verb, while the verb stem is required by the full form. It seems to be impossible to explain such alternations by a rule operating after lexical insertion. In addition, the meaning of kko (nai) is slightly different from the full form; the former is closer to V-Pres-ci hago (nai) 'it should not be the case that ... V ...' than to V-Pres-ci kko wa (nai) 'it is not the case that ... V ...'. If these forms are independent lexical items, the semantic differences between them is not problematic.

As for (13a), a lexical approach to the contracted forms, siman (Simaw) or Kaw (Jaw), is better than a phonological approach for the following reasons: (i) as specified in the rule, the process is peculiar to the fixed lexical items, namely an auxiliary element te (or de) and the perfect aspectual verb, siman 'finish'; (ii) although there seems to be a sequential change in this process, te (de) siman → siman (Simaw) → Kaw (Jaw), this sequence is not at all relevant to the rate of speech. The choice between the full form and a contracted form, either siman (Simaw) or Kaw (Jaw), is independent on how intimate the situation is. The choice between the two contracted forms seems to be idiosyncratic; for example, siman (Simaw) is not in my active vocabulary, while I use Kaw (Jaw) very often in intimate speech. Furthermore, in child speech in general, Kaw (Jaw) is used exclusively, and neither the full form nor the intermediate form seems to be in children's active vocabularies. Thus, the child seems to learn the most contracted form as an independent lexical item referring to the perfect aspect; (ii) the most contracted form, Kaw (Jaw), can appear in syntactic contexts, where the full form is not allowed. For instance, as Martin (1975) notices, Kaw (Jaw) can occur with a stative verb, such as ar 'be, exist', implying undesirability to the verb.

(17) a. kodomo ga ar-τaw-u to nani mo deki-nai.
   child SH exist-finish-pres. then everything even can-do-not
   I can't do anything because of having children.'
   b. *kodomo ga ar-te τaw-u to nani mo deki-nai.

If τaw (Jaw) is an independent item from the full form, the distribution difference in (17) is easily explained.

If the approach I am suggesting is correct, it will give consid-
erable difficulties to past analyses of the Japanese aspectual verb, *simaw*, where it is analyzed as a verb which takes either a sentential subject (18a) or a sentential complement (18b).

\[ (18) \]

\[ \begin{array}{c}
\text{S} \\
\text{V}
\end{array} \quad \begin{array}{c}
\text{simaw} \\
\text{te (de)}
\end{array} \]

To (18a), the cyclic transformation, Verb Raising (VR) applies to derive a surface structure, while to (18b) Equi NP Deletion, S-Pruning, and VR apply in this order. Regardless of the choice between (18a) and (18b), it is clear that the insertion of a contracted form, *simaw* (**simaw**) or *te (de)* (**te (de)**), by lexical insertion (at the level of deep structure) is impossible. Note that these contracted forms involve an Aux element in the embedded S and a verb in the matrix S. To accept the lexical analysis of contracted forms does not go along with the VR analysis of the aspectual verb, *simaw*. Rather it forces us to reanalyze the aspectual construction of Japanese in such a way that the contracted forms are lexically specified so as to take a verb phrase complement.

V. Summary

The terms 'fast' and 'casual', are interchangeably used in past literature on casual speech. In this paper, I attempt to make formal distinctions between FS and CS processes. To characterize such distinctions, I set up the criterion: whether a process is sensitive to the rate of speech, FS rules are such that their application is related to the rate of speech. CS processes are such that their operation is not depending on whether an utterance is produced fast or slowly. This criterion leads to further interesting differences between these two processes: (i) FS rules are not lexically governed, while CS rules are conditioned by lexical information; (ii) FS rules are neutral to the sociological factors. CS rules are, on the other hand, triggered by such factors.

In short, FS rules are operative purely on phonological grounds and their application is extended or restricted depending on how fast a speech is produced. Thus, they are best described in the phonological component. On the other hand, CS rules are sensitive to lexical information and sociological conditions. This is not only the characteristic of phonological processes of CS but that of morphological and syntactic processes of CS. To capture the generalization that CS processes are always conditioned by lexical information irrespective of components of grammar, it is suggested that all the CS phenomena should be described in the lexicon. This approach not only captures the relationship between lexicon and sociological conditioning in CS phenomena but also describes phenomena in a consistent fashion; exceptions, which must (if not all) CS processes possess, are no longer exceptions but follow from the lexical information given to each item. Furthermore, the lexical approach proposes a more restricted model of a grammar; sociological conditions are available only in the lexicon. Such a model is compatible with a current trend in syntax that strings generated by base rules are actually much more concrete than previously being thought (Bresnan (1978), Braine (1978)). In fact, the lexical approach to CS phenomena forces us to reanalyze certain syntactic configurations in a more concrete way (e.g. the analysis of Japanese aspectual verbs).

Footnotes

1. I would like to thank Ellen Kaisse for introducing me the interesting topic 'casual speech' and for valuable comments on an earlier version of this paper. Work on this study was supported by research grant AH-111 from the graduate school of the University of Washington, for which I am grateful.

2. I am using the term 'sociological', intending that it refers to notions like 'formal', 'informal', 'polite', 'intimate', 'familiar', 'vulgar', etc.

3. The feature '{pause} can be replaced by ', ', which stands for 'at the end of an intonational phrase', defined by Selkirk (1978).

4. In Hattori (1960), Kodaichi (1967), McCawley (1968), a glottal stop ' is posited before a vowel. Hence, the pairs I gave are written as su'i 'mirth principle', sur'i 'vessel vendor', satō'ya 'sugar salesman', and satō'yor 'foster parent'. If the existence of a glottal stop before a vowel which is not preceded by a consonant is well defended, the process I am describing can be taken as 'glottal stop deletion', which obeys the same condition as VDM.

5. There are several exceptions to the generalization concerning NS; the possessive no can be syllabified in front of certain lexical items (ia); no as a nominalizer can be syllabified before a question particle ka (1b); mono 'thing, object' can become mu (1c, 1d, le). Note that mono often appears in idiomatic expression (id, le). As you will see, these exceptional cases favor the lexical approach to CS processes, where they are no longer exceptions.

\[ (1) \]

\[ \begin{array}{c}
\text{buku no tokoro} \\
\text{boku} \quad \text{g tokoro} \quad \text{my place}
\end{array} \]

\[ \begin{array}{c}
\text{yatta no ka} \\
\text{yatta} \quad \text{g ka} \quad \text{(you) did (it)}
\end{array} \]

\[ \begin{array}{c}
\text{did VDM. Q}
\end{array} \]

\[ \begin{array}{c}
\text{soina mono} \\
\text{soina} \quad \text{mo (such a thing}
\end{array} \]

\[ \begin{array}{c}
\text{d. Širi mono ka} \\
\text{Širi} \quad \text{mo ka} \quad \text{How can (I) know?}
\end{array} \]

\[ \begin{array}{c}
\text{yatta mono da} \\
\text{yatta} \quad \text{mo da} \quad \text{used to do (it)}
\end{array} \]

The occurrence of ni as a copula is very restricted, namely it only appears before nar 'become' and su 'do'. ni as a particle (dative, directional) can occur before the copula, though it does not undergo NS.

\[ (1f) \]

\[ \begin{array}{c}
\text{tarō ni da} \\
\text{starō} \quad \text{g da} \quad \text{(it) is to Taro.}
\end{array} \]
6. Aspecual verbs, including *simax, have generally been analyzed as verbs which take a sentential subject or a sentential complement (see the section IV), rather than as inflectional affixes. When a sentence ends with *simax, it has to be followed by a negative predicate or a predicate denoting negation. This means the change is only possible when the higher predicate is negative, although the original string *simax + koto wa can be followed by an affirmative predicate. Thus, *simax + koto wa is a valid sentence.

7. How these forms are used and who uses them vary according to the situation, ideological factors, etc. Thus, what I give in (16) is not at all absolute or complete. The forms, a, b, c, d, e, f, and g in (16), are more or less phonologically related. Whatever the rules are to connect these forms phonologically, they are so lexically idiosyncratic that it is better to list these forms separately in the lexicon.

8. In Hattori’s model of the lexicon, refuse, for example, has the following lexical entry.

Features common to all derivatives including:
- NP
  - VP
  - [human]
  +V
  +Det

Features peculiar to the verb refuse
+ -al
(features indicating affixation of -al to form Nom.)
Other features peculiar to refuse

9. This result converges with that of Hasegawa (1978), where all the so-called dependent verbs (causative sense, aspeculal haleme begin, etc.) are analyzed as verbs with a verb phrase complement, resulting in a grammar without VR and other related ad hoc conventions and mechanisms.

10. Rotenberg (1978) independently reaches a similar conclusion. The FS rules discussed, HVD and VDBL, are what he calls “high-level phonological rules which are merely less punchy versions of independently necessary word-level ones.” (p. 200) He further characterizes such rules as follows: “the wider rule does the same thing as the narrow one, but applies with nothing like the same regularity, and only in connected speech. In every such case, I predict, there is in fact only one rule, an utterance level one sensitive to prosodic structure. Since prosodic structure changes with faster speech, generalizing itself over progressively larger domains, so will the applicability of the rule.” (pp. 200-201)

11. I regard English contraction processes (especially the contrac-