Journal of the Graduate School of Letters, Hokkaido University
Vol. 5; pp. 41–63, March 2010

The Past Tense Formation of Japanese Verbs: An Exploration in Cognitive Phonology

Masuhiro NOMURA

Abstract: Cognitive phonology, as proposed by Lakoff (1988, 1993) and Goldsmith (1989, 1990, 1991, 1993b), is a phonological theory that questions the foundational assumptions of generative phonology, doing away with the notions of "derivations" and "extrinsic rule ordering". The present paper is an attempt to explore possibilities and problems of cognitive phonology by analyzing the past tense formation of Japanese verbs as a case study. It is shown that the framework of cognitive phonology nicely analyzes the data, but at the same time such problems as relationships across levels, environment statements, and measurement of harmony are pointed out and discussed.

(Received on December 18, 2009)

1. Introduction

With the recent development of cognitive science and linguistics, a number of new phonological theories have been proposed that challenge such fundamental assumptions of classical generative phonology as derivations, extrinsic rule ordering, cycles, and application principles. Among them are "cognitive phonology" proposed by Lakoff (1988, 1993) and "harmonic phonology" proposed by Goldsmith (1989, 1990, 1991, 1993b). Since these two share fundamental assumptions and a theoretical framework, I will use the term "cognitive phonology" in this paper to cover both. The present paper is an attempt to describe the past tense formation of Japanese verbs in the framework of cognitive phonology and, in so doing, consider possibilities and problems of cognitive phonology.[1]

The organization of the paper is as follows: Section 2 outlines the basic assumptions and claims of cognitive phonology, focusing especially on how they differ from those of classical generative phonology. Section 3 describes the phonotactics that are at work in Japanese.

1 This paper is a slightly modified version of the comprehensive paper I submitted to the Department of Linguistics, University of California at San Diego, in May 1994. The framework adopted in the present paper has not, unfortunately, gained much ground in subsequent developments of cognitive linguistics; for recent accounts of cognitive phonology, see Kumashiro (2000), Bybee (2001) and Nathan (2008).

Masuhiro NOMURA: nomura@let.hokudai.ac.jp

Section 4 proposes a cognitive phonology analysis of the past tense formation of Japanese verbs. Section 5 discusses advantages and disadvantages of the proposed analysis. Section 6 is devoted to conclusion.

2. Outline of Cognitive Phonology

2.1 The Basics

It is reasonable to say that the conception of cognitive phonology has been motivated by careful reconsideration of the relationship between linguistic levels and derivations, as indicated by the following quote from Goldsmith (1993b):

(1) Inherent in the notion of level is the idea that there are specific generalizations that can, and perhaps must, be drawn concerning the representation of the expression on that level. For as we said earlier, a level is a way of representing, or describing, or analyzing, an utterance; many (perhaps a boundless number of) levels are thus conceivable, and the correct ones will be justified on the basis of the generalizations that can be stated at just those levels. (Goldsmith 1993b: 27–28)

Goldsmith suggests that the notion of level is best explicated in traditional structuralist phonology, where three quasi-phonological representations were posited: a phonetic representation (PT), a phonemic representation (PM) and a morphophonemic representation (MP). These three representations enable us to describe phonemic alternation rules and allophony rules as (MP, PM) rules and (PM, PT) rules, respectively, thereby assuring an inherent ordering between the rules (Goldsmith 1993b: 26).

By contrast, classical generative phonology dismissed a phonemic level from phonological representations, as a consequence of which "Ordering, quite necessary in linguistic accounts as we have seen, became then (on this generative account) not the function of relations across levels, but a property inherent to another type of object, a rule, whose function was to create something that was not a representation on any particular linguistic level. And that was a very special innovation in linguistic theory: a representation that was of no particular linguistic level. Such representations are the intermediate forms of generative grammar" (Goldsmith 1993b: 27).[2]

Cognitive phonology assumes that a level L consists of (2) and posits three levels of phonological representation as a minimal collection of necessary dimensions (M, W, and P stand for "morpheme", "word", and "phonetic", respectively), as summarized in (3):[3]

2 See Schane (1971) for an argument that phonemes do have a position (though not occupying "a level of representation") in generative phonology.

3 The definitions in (3) are not necessarily crystal-clear, but it may be assumed that "M-level" and "P-level" roughly correspond respectively to "underlying representation" and "phonetic representation" in the traditional sense, with the difference that each level in cognitive phonology is a series of representations where the last one best satisfies the level's phonotactics. In Goldsmith (1990: 330), M-level is defined as "the level at which *segmentally* represented morphemes are represented" (italics mine) and "Its sole function is as a repository of the minimal information necessary to capture the sound characteristics of the morpheme." Lakoff (1993), on the other hand, uses distinctive features to represent M-level in some of his concrete analyses.

(2) a. a vocabulary permitting a linguistic description[4]
b. a set of relations expressing relative well-formedness
c. a set of intralevel (L, L) rules which express the paths that a representation may pass through to find maximal well-formedness (Goldsmith 1991: 253–257, 1993b: 31)

(3) • *M-level*, a morphophonemic level, the level at which morphemes are phonologically specified
• *W-level*, the level at which expressions are structured into well-formed syllables and well-formed words, but with a minimum of redundant phonological information
• *P-level*, a level of broad phonetic description that is the interface with the peripheral articulatory and acoustic devices. (Goldsmith 1993b: 32)

By positing these three levels, cognitive phonology proposes to replace the functions of derivations by two types of relations: intralevel constructions, which state well-formedness conditions within levels, and cross-level (or interlevel) constructions, which state correlations across levels.[5] These two types of constructions differ with respect to the way they apply (Goldsmith 1993b: 37):

(4) a. intralevel constructions: (M, M), (W, W), and (P, P)
• harmonic in function[6]
• free reapplication
b. cross-level constructions: (M, W) and (W, P)
• may or may not be harmonic in function
• simultaneous application[7]

The overall model of cognitive phonology is summarized below (Goldsmith 1990: 324, 1991: 257, 1993b: 33):

4 Goldsmith is not clear about what "vocabulary" is specifically necessary for each level, simply stating that "with regard to matters of phonological representation, at least, this harmonic phonology will need all the mechanisms that recent phonological theory has offered" (Goldsmith 1993b: 30). See also Wiltshire (1992: 4).

5 To express these two types of relations, Lakoff (1988, 1993) uses the term "constructions", while Goldsmith (1993b) prefers to use the traditional term "rules". I will use the terms "constructions" and "rules" interchangeably in this paper.

6 Rule application is "harmonic" when "they (=rules) apply just in case their output is better than their input with respect to some criteria specified by a phonotactic (of the relevant level)" (Goldsmith 1991: 252). Lakoff seems to hold a different position from Goldsmith on "harmonic" application when he remarks that "I conceive of constructions as increasing harmony (in Smolensky's sense) *both within and across dimensions*" (Lakoff 1993: 141; italics mine).

7 "Constructions combine by superposition. That is, each construction imposes a set of constraints, and the constraints of the various constructions are simultaneously satisfied" (Lakoff 1993: 118). See also (Lakoff 1988: 306).

(5) M (M, M) intralevel (harmonic)

↕ ←(M, W) cross-level (harmonic or nonharmonic)

W (W, W) intralevel (harmonic)

↕ ←(W, P) cross-level (harmonic or nonharmonic)

P (P, P) intralevel (harmonic)

Goldsmith (1990: 324) gives the following examples as representing each intralevel and cross-level construction:[8]

(6)

Type of rule	Example
(a) (M, M)	Melody spreading before tier conflation
(b) (M, W)	Tier conflation
(c) (W, W)	Syllabification; epenthesis
(d) (W, P)	Default feature specification
(e) (P, P)	Flap formation in English

2.2 Rule Interactions in Cognitive Phonology

There are a couple of points worth mentioning with respect to the model in (5). First, because of free (re)applicability of intralevel constructions, a level does not constitute a single representation, but a series of representations where the last one best satisfies the level's phonotactics. Second, there is no directionality among the three levels: "There is no object moving or being translated from one of the three levels to another; what is 'happening' during the linguistic analysis is that three (simultaneous) types of analysis are being compared, contrasted, and measured for fit" (Goldsmith 1993b: 45).[9]

The above two points should be understood in the context of PDP (Parallel Distributed Processing) connectionism, which offers a new perspective on our cognitive processing (Rumelhart and McClelland 1986). The dominant theoretical paradigm of cognitive science has long been shaped by the so-called "von Neumann computer metaphor", according to which a single central processing unit executes a series of a complex algorithmic program (cf. Langacker 1987b). PDP connectionism replaces this view with the "neurally inspired" view that a large set of units carries out simple arithmetic computations in a parallel fashion. Each unit is connected

8 See Goldsmith's (1991, 1993b) analysis of Lardil for illustration of the difference between intralevel constructions (harmonic, free (re)applicability) and cross-level constructions (which may or may not be harmonic). As Schane (1971) demonstrates, the same rule could serve a morphophonemic function in one language but a phonetic function in another, which means, in cognitive phonology terminology, that the same construction could function as a (M, W) construction in one language, and a (W, P) construction in another.

9 Lakoff (1988: 308) states that "multiple activation patterns can be activated simultaneously in real time in real neural networks. Phonological constructions should be representable in terms of such patterns".

to other units with a certain connection weight and computes whether it should be on or off, based on the incoming activation and the connection weight. A "representation" is a state of a network, which is expressed as the on/off state of each of its units. Goldsmith adopts a network system where units continue to recompute until the continued computation no longer has observable effects, i. e., until the network settles into a state of equilibrium that maximizes the harmony (cf. Smolensky 1986: 210–213, Goldsmith 1993a: 10). In other words, a "representation" shifts toward a state of equilibrium, which constitutes the final and best-formed representation on a given level. Seen from this connectionist perspective, the "shift" of representations should be conceptually distinguished from "derivation" in the generative tradition.[10]

We are now in a position to see how the notion of "derivation" in classical generative phonology is handled by cognitive phonology. Classical generative phonology adopts the ordered-rule hypothesis (i. e. phonological rules apply in sequence and the sequencing of rules is (at least partially) determined by ordering statements) and describes kinds of rule interactions as follows:[11]

(7) a. If R_i counterfeeds R_j, then R_j must precede R_i
b. If R_i counterbleeds R_j, then R_j must precede R_i
c. If R_i feeds R_j, then R_i must precede R_j
d. If R_i bleeds R_j, then R_i must precede R_j

Rule interactions are common in natural languages and cognitive phonology does not deny their existence (cf. Goldsmith 1993a: 6). Cognitive phonology, however, finds derivation with an ordered set of rules problematic from the viewpoint of learnability and processing (see Goldsmith 1993a: 5–6, 1993b: 54–55; Lakoff 1988: 308, 1993: 117), taking the position that "within the phonological part of a grammar, just as within the other components of the grammar, the acquisition of a language consists of the abstraction of a large number of well-formedness conditions — patterns, crudely put — on a small number of levels — three" (Goldsmith 1993b: 30).[12]

10 Goldsmith (1993b: 31) admits that "Even this model [=(5)] is unsatisfactorily derivationally oriented", stating that "the difficulty in avoiding even these remnants of derivationalism lies in the question of how to compose two or more distinct phonological rules; without a more radical revision of the notion of a phonological representation, it is difficult, and perhaps impossible, to treat the effects of several rules without some remnants of derivationalism, which is to say, a linear sequence of distinct, identifiable representations". See Goldsmith (1993b: 46–56) for his Dynamic Computational Models, which is an attempt to overcome this difficulty.

11 Goldsmith (1991: 249, 1993b: 28–29) suggests that the notion of derivation has two sources: one is logicians' formalization of logical derivation and the other is regular sound change in historical linguistics. He then states that "it is far too easy for us, in our present position, to think of derivations in phonology as arising somehow, ineluctably and logically, jointly out of the data and the task. It is, rather, in large measure a historical, or even biographical, matter that the theoreticians who have most influenced our current views on this (of whom the first and foremost is Chomsky) have offered us this particular view; others are equally congenial to the task" (Goldsmith 1993b: 29).

12 This view is basically compatible with Langacker's (1988) usage-based model, though Langacker is not committed to the number of levels. See also Langacker (1987a: 442–445) for his cognitive grammar analysis of feeding relationships.

As mentioned in the previous section, cognitive phonology proposes to replace the notion of derivations by cross-level constructions and intralevel constructions. Counterfeeding/counterbleeding interactions are handled by applying two constructions in the same cross-level component (i. e., either (M, W) or (W, P)), so that they will not interact.[13] Feeding/bleeding interactions are typically handled by applying two constructions in different cross-level components (i. e., (M, W) and (W, P)) or applying one construction in a cross-level component and the other in an intralevel component (e. g. (M, W) and (W, W)).[14]

3. Phonotactics in Japanese

As a preliminary to our analysis of the past tense formation of Japanese verbs, this section discusses well-formedness conditions that hold at level M and level W in Japanese.

3.1 M-level Phonotactics

There is a morpheme structure constraint on Yamato (native) Japanese vocabulary that says that a morpheme may contain at most one voiced obstruent (Itô and Mester 1986: 67):

(8) $*[[+\text{voi}]\ [+\text{voi}]]_{\text{morpheme}}$

This constraint allows a morpheme like *buta*, but disallows a morpheme like **buda*.

Itô and Mester (1986: 68) suggest that this morpheme structure constraint falls automatically out of Lyman's Law. Lyman's Law deletes a [+voi] specification followed by another [+voi] specification on the voicing tier:

(9) Lyman's Law (Itô and Mester 1986: 60)

$$[+\text{voi}] \rightarrow \emptyset \ / \ \underset{\substack{| \\ x'}}{\_\_}\ [+\text{voi}]$$

(x' marks a skeletal slot)

Lyman's Law is relevant to the analysis of *rendaku* (sequential voicing), in which an initial obstruent of the second component of a compound noun gets voiced:[15]

13 See Lakoff's (1993) formulations of Vowel Shortening and Vowel Lowering in Yawelmani, and Mid Vowel Raising and Vowel Assimilation in Baztan. Note that this treatment of counterfeeding and counterbleeding interactions is essentially a resurrection of the "direct mapping hypothesis" (cf. Kenstowicz and Kisseberth 1979: Ch. 8), which classical generative phonology rejected. Needless to say, cognitive phonology is different from the direct mapping hypothesis, most notably, in its postulation of three levels (instead of two).

14 See Lakoff's (1993) formulations of Truncation and Vowel Shortening in Yawelmani (feeding), Epenthesis and Vowel Shortening in Yawelmani (bleeding), and Goldsmith's (1993b) formulations of Apocope and Non-apical Consonant Deletion in Lardil (feeding).

15 Itô and Mester (1986: 60) analyze *rendaku* as "a [+voi] autosegment linked to a skeletal slot and associated with its surface bearer by the general rule of Voicing Spread".

(10) a. iro ('color') + **k**ami ('paper') → iro**g**ami ('colored paper')
b. yo ('night') + **s**akura ('cherry') → yo**z**akura ('blossoms at night')
c. e ('picture') + **t**ako ('kite') → e**d**ako ('picture kite')
d. ike ('arrange') + **h**ana → ike**b**ana ('ikebana')[16]

Rendaku is blocked by Lyman's Law when the second component of a compound contains a voiced obstruent:

(11) a. kami + ka**z**e → *kami**g**a**z**e ('divine wind')
b. mono + si**z**uka → *mono**z**i**z**uka ('tranquil')
c. siro + ta**b**i → *siro**d**a**b**i ('white tabi')
d. maru + ha**d**aka → *maru**b**a**d**aka ('completely naked')

Itô and Mester (1986: 68) suggest that "let Lyman's Law apply within morphemes. Since it deletes a [+voi] in the context of another [+voi], it already ensures that only one [+voi] can appear within a morpheme" and that "The morpheme structure constraint [=(8)], though expressing a true generalization, is a mere corollary of the phonological rule of Lyman's Law and thus has no independent theoretical status."

By contrast, since a conspiracy effect is attributed, in cognitive phonology, to phonotactics rather than to phonological rules, I would suggest that constraint (8) serves as an M-level phonotactic. In addition, I suggest that this constraint is also operative at level W, to prevent *rendaku* from applying to the examples in (11):

(8)' $*[[+\text{voi}]\ [+\text{voi}]]_{\text{word}}$

3.2 W-level Phonotactics

3.2.1 Syllable Structure in Japanese

Japanese basically allows the following four types of syllable patterns (Itô 1986: 18):[17]

(12) a. CV (e. g. <u>ka</u>•mi•kaze 'divine wind')
b. CVV (e. g. <u>kai</u>•soo 'seaweed')
c. CVN (e. g. <u>kam</u>•pai 'cheers')
d. CVC (e. g. <u>sek</u>•ken 'soap', <u>kap</u>•pa 'legendary being', <u>tos</u>•sa 'impulsively')

Itô (1986: 26) proposes the following Coda Condition for Japanese:[18]

16 The alternation /h/ ~ /b/ is the result of a historical change where /p/ changed into /h/.

17 In addition, superheavy syllables are possible, though they are more marked than those in (12): CVVC (e. g. <u>toot</u>•te), CVNC (e. g. ron•<u>donk</u>•ko).

18 This is equivalent to saying that the coda is incapable of licensing a point of articulation autosegment: "nasals and obstruents that appear in the coda will either be homorganic to a following consonant — that is, they will share the point of articulation autosegment that is licensed by the following onset — or they will have a non-distinctive default point of articulation" (Goldsmith 1990: 125).

```
(13)   * C]σ
          |
       [-nas]
```

On the assumption that Hayes' (1986: 331) "Linking Constraint" (i. e. Association lines in structural descriptions are interpreted as exhaustive) is applicable to the interpretation of phonotactic constraints, Coda Condition (13) allows (14a) and (14b), but not (14c):

```
(14) a. hon          b. kappa              c. *kapta
          σ               σ      σ              σ     σ
        / | \           / | \   / \           / | \  / \
       C  V  C         C  V  C  C  V         C  V  C  C  V
       |  |  |         |  |   \/   |         |  |  |  |  |
       h  o  n         k  a   p    a         k  a  p  t  a
```

Since /p/ in the coda of (14b) is doubly-linked, it does not violate Coda Condition (13) under the Linking Constraint interpretation. By contrast, /p/ in the coda of (14c) is singly-linked, thereby violating (13) to make the syllable impermissible.

3.2.2 Permissible Consonant Clusters

In addition to Coda Condition (13), the Japanese language is subject to the following constraints on consonant clusters (cf. Itô and Mester 1993):

(15) *DD Constraint

Voiced obstruent geminates are not allowed; only voiceless obstruent geminates are allowed (*bb, *dd, *gg, *zz; pp, tt, kk, ss).

```
*C  •  C
  \   /
[+voice]
```

(16) *NT Constraint

A nasal cannot be followed by a voiceless obstruent (*mp, *nt, *ŋk, *ns; mb, nd, ŋg, nz).

```
*N     C
    [-voice]
```

4. The Past Tense Formation of Japanese Verbs

This section aims to apply the framework of cognitive phonology to Japanese verb conjugations involving the past tense suffix *-ta*.

4.1 Data and Generalizations

The data to be analyzed in this section are listed below, where the left-hand column represents the root form of a verb and the right-hand column represents its past tense form:[19]

(17)	a.	tabe-	'to eat'	tabeta	'ate'
	b.	mi-	'to see'	mita	'saw'
(18)	a.	tor-	'to take'	totta	'took'
	b.	kaw-	'to buy'	katta	'bought'
(19)	a.	sin-	'to die'	sinda	'died'
	b.	yom-	'to read'	yonda	'read'
	c.	tob-	'to fly'	tonda	'flew'
(20)	a.	kak-	'to write'	kaita	'wrote'
	b.	kag-	'to sniff'	kaida	'sniffed'
	c.	kas-	'to rent'	kasita	'rented'

The verbs in (17) are vowel-ending verbs and those in (18)–(20) are consonant-ending verbs. The past tense formation of the former is quite straightforward: the past tense suffix *-ta* is simply added to the root. By contrast, various alternations are evidenced in (18)–(20). In (18), the root-final non-nasal sonorants /r/ and /w/ are totally assimilated to /t/. In (19), the past tense suffix *-ta* changes to *-da*, while the root-final consonants /m/ and /b/ change to /n/ in front of the past tense suffix. In (20a, b), the root-final velar is deleted and the vowel /i/ is epenthesized (additionally, *-ta* changes to *-da* in (20b)). In (20c), the vowel /i/ is simply epenthesized, while the preceding /s/ remains intact.

The past tense formation outlined above can be seen as processes of syllabifying M-level consonant clusters /rt/, /wt/, /nt/, /mt/, /bt/, /kt/, /gt/, /st/, which violate the W-level well-formedness conditions.

4.2 Constructions

The basic data having been presented, we can begin to posit constructions that apply to shift the representations toward the state that maximally satisfies the phonotactics at level W (and level P).

4.2.1 Gemination

Gemination is a well-attested means of resolving unsyllabifiable consonant clusters, as observed in compound verb formation (21), Sino-Japanese compound words (22), loanwords (23), and casual/fast speech (24):

(21) a. fuk ('blow') + tobas ('fly') → futtobas ('blow away')
hik ('pull') + tigir ('tear') → hittigir ('tear off')

---

19 For previous studies of this topic, see Ashworth (1976–77), Davis and Tsujimura (1991), Itô and Mester (1986), Kuroda (1965, 1966), McCawley (1968), Poser (1986), Yoshiba (1981) and Zec (1993).

hik ('pull') + ture ('take along') → hitturer ('take along')
b. but ('hit') + kom ('be full') → bukkom ('throw into')
c. nor ('ride') + kir ('cut') → nokkir ('get over')
d. ow ('chase') + tuk ('arrive') → ottuk ('overtake') (Poser 1986: 172–3)

(22) a. kot + kaku → kokkaku ('frame')
b. sit + pai → sippai ('failure')
c. rit + syoo → rissyoo ('proof')
d. it + sai → issai ('one year old')
e. rok + pai → roppai ('six cupfuls')
(Kuroda 1965: 207, McCawley 1967: 85, Itô 1986: 144–154, Vance1987: 138–160)

(23) Russian /votka/ > uokka (Poser 1986: 173)

(24) wakaranai ~ wakannai 'does not understand' (Poser 1986: 173)

In cognitive phonology, gemination can be formulated as an M-W construction:

(25) Gemination
M: C_i
|
W: C_j C_j

It follows from the W-level phonotactic in (13) that C_i is [-nas], which is unsyllabifiable as a coda at level W, and that there is a syllable boundary between two C_j's at level W.[20]

This construction sanctions the following analysis of the verbs in (18):

(26) a. M: tor + ta
(Gemination)
W: totta
b. M: kaw + ta
(Gemination)
W: katta

4.2.2 Voiced Geminate Nasalization

Gemination (25) could potentially apply non-harmonically if only a sequence of consonants meets its environment conditions. Thus, when C_j is a voiced obstruent, a voiced obstruent geminate (bb, dd, gg, zz) will result, in violation of the W-level phonotactic of *DD Constraint (15). In case this happens, a W-level intralevel construction must apply, as a repair strategy, to remedy the violation (recall that all intralevel constructions are harmonic in function). This intralevel construction nasalizes the first component of an impermissible voiced obstruent geminate *DD, so that it may satisfy both Coda Condition (13) and *NT Constraint (16). This intralevel construction is formulated as follows:

20 The reason the environment condition is stated at level W will be discussed in Section 5.2.

(27) Voiced Geminate Nasalization[21]

W: $D_i \rightarrow N / \_\_ D_i$ (D = voiced obstruent, N = mora nasal)

It follows from (12) and (13) that any sequence of two consonants is separated by a syllable boundary. Thus, the two D_i's in (27) are separated by a syllable boundary. More visually, (27) can be represented as (27)':

(27)' W: [+nas]

↓

C_i $ C_i

＼ ／

[+voice]

This construction is independently motivated by compound verb formation in (28) and intensive infixation in (29):

(28) a. tuk ('stab') + das ('put out') → tundas ('thrust out')

b. ow ('chase') + das ('put out') → ondas ('drive out') (Poser 1986: 171–173)

cf. ow ('chase') + tuk ('arrive') → ottuk ('overtake')

(29) koNgari, maNziri, noNbiri

cf. haQkiri, baQsari, baQtari, suQpori[22]

(Kuroda 1965: 205–206, 1966: 90, McCawley 1968: 97, Itô and Mester 1986: 59)

Notice that in both compound verb formation and intensive infixation, the mora nasal appears when followed by a voiced consonant, while the mora obstruent appears when followed by a voiceless consonant.

Gemination (25) and Voiced Geminate Nasalization (27) sanction the following analysis of the examples in (28) and (29):

(30) M: ow + das

(Gemination)

W: oddas

(Voiced Geminate Nasalization)

oNdas

21 This construction is essentially equivalent to Coda Nasalization (Itô and Mester 1986) and Voiced Geminate Nasalization (Poser 1986).

22 Q and N represent the "mora obstruent" and the "mora nasal", respectively. See Vance (1987: Ch. 5) and Kuroda (1966) for details.

(31) M: noCbiri (C = intensive infix)
(Gemination)
W: nobbiri
(Voiced Geminate Nasalization)
noNbiri

4.2.3 Voicing Assimilation

Let us look once again at (18)–(20), repeated below, paying attention to the past tense suffix alternation between *-ta* and *-da*:

(18)	a.	tor-	'to take'	totta	'took'
	b.	kaw-	'to buy'	katta	'bought'
(19)	a.	sin-	'to die'	sinda	'died'
	b.	yom-	'to read'	yonda	'read'
	c.	tob-	'to fly'	tonda	'flew'
(20)	a.	kak-	'to write'	kaita	'wrote'
	b.	kag-	'to sniff'	kaida	'sniffed'
	c.	kas-	'to rent'	kasita	'rented'

Comparison between (20a) and (20b) reveals that the contrast between *-ta* and *-da* is attributable to the voicing contrast of the stem-final velar consonants /k/ and /g/. This is also true of (19c), where voicing is observed after /b/-stems. We can formulate this observation as an M-W construction that states that when a voiced segment is adjacent to a voiceless segment at level M, the latter becomes voiced at level W:[23]

(32) Voicing Assimilation
M: [+voice] [-voice]
|
W: [+voice]

Notice that this construction applies non-harmonically, creating two voiced obstruents at level W, which is in violation of the W-level phonotactic (8)'.

Voicing Assimilation, together with Gemination and Voiced Geminate Nasalization, sanction the following analysis of the past tense formation of *tob-* 'to fly':

(33) M: tob + ta
(Voicing Assimilation, Gemination)
W: todda
(Voiced Geminate Nasalization)
toNda

23 This construction is essentially equivalent to Voicing Assimilation (McCawley 1968, Poser 1986, Davis and Tsujimura 1991) and Voicing Spread (Itô and Mester 1986).

4.2.4 Mora Nasal Formation

Given that a coda position is incapable of licensing a point of articulation autosegment (cf. (13)), I suggest that a nasal segment followed by a consonant at level M becomes the mora nasal N, unspecified for place, when syllabified at level W. This can be formulated as follows:

(34) Mora Nasal Formation

M: [+nas] C
        |
W: N

This construction sanctions the following analysis of the past tense formation of *yom* 'to read':

(35) M: yom + ta

(Mora Nasal Formation)

W: yoNta

4.2.5 Postnasal Voicing

A nasal followed by a voiceless obstruent at level W violates *NT constraint described in (16). When this happens, an intralevel construction must apply to shift the representation to allow maximal satisfaction of the W-level phonotactics. This intralevel construction can be formulated as follows:

(36) Postnasal Voicing

W: [-son, -voice] → [+voice] / N ____

This construction, together with Mora Nasal Formation, shifts the representation in (35) as follows:

(37) M: yom + ta

(Mora Nasal Formation)

W: yoNta

(Postnasal Voicing)

yoNda

Voicing Assimilation and Postnasal Voicing have in common that they trigger a change from [-voice] to [+voice]; they differ, however, in that the former is a cross-level construction that could apply non-harmonically, while the latter is an intralevel construction, applying harmonically as a repair strategy.

4.2.6 Nasal Assimilation

The next construction is Nasal Assimilation, which assimilates the place of articulation of the mora nasal N to that of the following consonant. Nasal Assimilation takes place not only across a morpheme boundary (38), but also across a word boundary (39):

(38) a. hoN ('book') + dana ('shelf') > ho**n**dana ('bookshelf')
b. hoN ('book') + bako ('box') > ho**m**bako ('bookcase')
c. hoN ('book') + gumi ('set, compose') > ho**ŋ**gumi ('makeup')
(39) a. hoN ('book') da ('copula') > ho**n** da ('is a book')
b. hoN ('book') bakari ('only') > ho**m** bakari ('book only')
c. hoN ('book') ga ('NOM marker') > ho**ŋ** ga ('book is')

Since Nasal Assimilation applies without regard to the morpheme constituency of the phonological string, I will formulate this construction as a W-P construction, on the assumption that there is a P-level well-formedness condition that says "In the sequence NC, the two consonants must share a place of articulation":

(40) Nasal Assimilation
W: $N_{\alpha place}$ $C_{\beta place}$
|
P: $N_{\beta place}$

Now, this construction, in concert with Mora Nasal Formation and Postnasal Voicing, completes the analysis of *yom* + *ta* ('read') in the following way:

(41) M: yom + ta
(Mora Nasal Formation)
W: yoNta
(Postnasal Voicing)
yoNda
(Nasal Assimilation)
P: yonda

4.2.7 Epenthesis 1

Epenthesis is a well-attested means to resolve unsyllabifiable consonant clusters:

(42) a. hik ('pull') + mekur ('strip off') → hikimekur ('peel')
b. but ('hit') + kom ('be full') → butikom ('throw into')
c. nor ('ride') + kir ('cut') → norikir ('get over')
d. ow ('chase') + tuk ('arrive') → oituk ('overtake')[24]
e. ow ('chase') + das ('put out') → oidas ('drive out') (Poser 1986: 172)

24 Present-day Japanese has a rule: w → Ø / __ i, u, e, o. This rule is responsible for the deletion of /w/ in (42d, e).

In order to explain (20), repeated below, two types of epenthesis are necessary, Epenthesis 1 for velar-final verb stems, and Epenthesis 2 for /s/-final verb stems:[25]

(20) a. kak- 'to write' kaita 'wrote'
b. kag- 'to sniff' kaida 'sniffed'
c. kas- 'to rent' kasita 'rented'

Epenthesis 1 deletes a stem-final velar and insert /i/, thereby changing the unsyllabifiable clusters /kt/ and /gt/ into the permissible CVV•CV pattern (e. g. kai•ta).[26] As the examples in (21a), repeated below, indicate, gemination is yet another means available to resolve these unsyllabifiable clusters:

(21a) fuk ('blow') + tobas ('fly') → futtobas ('blow away')
hik ('pull') + tigir ('tear') → hittigir ('tear off')
hik ('pull') + ture ('take along') → hitturer ('take along') (Poser 1986: 173)

This justifies us in formulating Epenthesis 1 as a morphologically conditioned construction where the past tense suffix *-ta* is explicitly mentioned at level M:[27]

(43) Epenthesis 1

M:	C	+	CV
	[velar]		ta
	|	|	
	x	x	
	|	|	
W:		i	

25 By contrast, Davis and Tsujimura (1991) posit a single epenthesis rule for the two types of verb stems and a velar deletion rule for the velar-final verb stems. We could capture their idea in cognitive phonology by positing Generalized Epenthesis as an M-W construction and Velar Deletion as a W-level intralevel construction. The problem is that the latter is not particularly harmonic, because *kakita* is as well-formed as *kaita*, both observing the W-level phonotactics in Japanese. The two epenthesis constructions reflect historical changes called in Japanese philology *i-onbin*, in which /ki/, /gi/ and /si/ changed to /i/ during the Heian Period (9th –12th century): e. g. *kakite* > *kaite* ('write'), *kagite* > *kaide* ('sniff'), *otosite* > *otoite* ('drop'). The deleted /s/ somehow revived later in the history of then Standard Japanese, which accounts for the two distinct types of epenthesis constructions. Thus, the difference in the past tense formation between velar-final verb stems and /s/-final verb stems is merely accidental and should not be attributed to the phonetic/phonological differences between /k, g/ and /s/.

26 Anderson and de Chene (1979) and Itô and Mester (1986) argue that the velar in question may turn into /i/ by "Velar Gliding (Vocalization)". This analysis is criticized by Poser (1986), who demonstrates that velar-glide alternation is morphologically governed, rather than a purely phonological rule.

27 cf. "One of the attractions of cognitive phonology is that, since it is part of cognitive grammar, other aspects of the grammar are directly accessible. Since cognitive grammars characterize correlations across various dimensions of structure, correlations between the phonology and various aspects of syntax, semantics, and pragmatics are directly statable" (Lakoff 1993: 144).

Since Epenthesis 1 is more specific than Gemination, the former preempts the latter by the Elsewhere Condition (Kiparsky 1973).[28] Thus, the following analysis is sanctioned:

(44) a. M: kak + ta
(Epenthesis 1)
W: kaita

b. M: kag + ta
(Epenthesis 1, Voicing Assimilation)
W: kaida

4.2.8 Epenthesis 2

The other type of epenthesis, Epenthesis 2, simply inserts /i/ between the stem-final /s/ and the suffix *-ta*, thereby changing the unsyllabifiable cluster /st/ to the permissible CV•CV pattern at level W, as in:

(45) kas- ('to rent') + ta → ka•si•ta ('rented')

Since gemination is yet another means available to resolve the unsyllabifiable cluster /st/ (e. g. o**s** ('push') + **t**ubus ('crush') → o**tt**ubus ('flatten')), Epenthesis 2 is a morphologically conditioned construction, just like Epenthesis 1, and can be formulated as an M-W construction:

(46) Epenthesis 2
M: C + CV
s ta
|
x
|
W: i

This construction sanctions the following analysis:

(47) M: kas + ta
(Epenthesis 2)
W: kasita

28 The Elsewhere Condition can be viewed as an automatic consequence of PDP connectionism, according to which "Generalizations involve low activation levels over large regions of the network; special cases involve higher activations over small regions of the network, and these override the low activations of the general cases" (Lakoff 1988: 306). See also Lakoff (1993: 145).

4.3 Summary of Analysis

We posited eight constructions above to account for the past tense formation of Japanese verbs:

(25) Gemination

M: C_i
|
W: C_j C_j

(27) Voiced Geminate Nasalization

W: $D_i \rightarrow N/ \_\_ D_i$ (D = voiced obstruent, N = mora nasal)

(32) Voicing Assimilation

M: [+voice] [-voice]
|
W: [+voice]

(34) Mora Nasal Formation

M: [+nas] C
|
W: N

(36) Postnasal Voicing

W: [-son, -voice] —> [+voice] / N ____

(40) Nasal Assimilation

W: $N_{\alpha place}$ $C_{\beta place}$
|
P: $N_{\beta place}$

(43) Epenthesis 1

M: C + CV
[velar] ta
| |
x x
| |
W: i

(46) Epenthesis 2

M: C + CV
s ta
|
x
|
W: i

The above set of constructions sanction the following analyses of the past tense formation of Japanese verbs listed in (18)-(20):

(18a)' M: tor + ta
(Gemination)
W: totta
P: totta

(18b)' M: kaw + ta
(Gemination)
W: katta
P: katta

(19a)' M: sin + ta
(Mora Nasal Formation)
W: siNta
(Postnasal Voicing)
siNda
(Nasal Assimilation)
P: sinda

(19b)' M: yom + ta
(Mora Nasal Formation)
W: yoNta
(Postnasal Voicing)
yoNda
(Nasal Assimilation)
P: yonda

(19c)' M: tob + ta
(Voicing Assimilation, Gemination)[29]
W: todda
(Voiced Geminate Nasalization)
toNda
(Nasal Assimilation)
P: tonda

(20a)' M: kak + ta
(Epenthesis 1)
W: kaita
P: kaita

29 Voicing Assimilation and Gemination apply simultaneously at the M-W component.

(20b)' M: kag + ta

(Epenthesis 1, Voicing Assimilation)[30]

W: kaida

P: kaida

(20c)' M: kas + ta

(Epenthesis 2)

W: kasita

P: kasita

5. Discussion

In this section, I will discuss advantages and disadvantages of the analysis proposed in the preceding section.

5.1 Advantages

First, the proposed analysis eliminates a long set of ordered rules, most clearly assumed by Davis and Tsujimura (1991). To begin with, the ordering between Epenthesis 1 and Voicing Assimilation (to distinguish between *kak + ta* → *kaita* and *kag + ta* → *kaida*) is resolved by the notion of simultaneous application, which enables these two constructions to apply simultaneously in the M-W component, as shown in (20b)'. Next, the Elsewhere Condition ensures that Epenthesis 1 and Epenthesis 2, on the one hand, and Gemination, on the other, do not have to interact, since the former are more special than the latter. Thus, our analysis need not stipulate, by rule ordering, that Gemination not change, for instance, *kas + ta* into *katta*. Similarly, owing to the Elsewhere Condition, the distinction between *tob + ta* → *tonda* and *kag + ta* → *kaida* is maintained without any stipulation: there is simply no possibility in this framework that *kag + ta* may become *kanda*, in the same way that *tob + ta* becomes *tonda*.

Second, the proposed analysis accounts for all the data by a set of constructions that are motivated on independent grounds: the same constructions are necessary to account for verb compound formation, intensive infixation, and so on.

Third, the construction-based approach of cognitive phonology can deal with the epenthesis cases (cf. (20), (20)'), which are morphologically conditioned, as we saw earlier. Formulating constructions to capture these morphologized epenthesis cases is unproblematic in cognitive phonology, since a construction serves to characterize well-established and entrenched cross-dimensional correlations.

5.2 Disadvantages

Several problems can be identified for the proposed analysis and some of them may well be related to the framework of cognitive phonology itself.

The first problem concerns relationships across levels. In Goldsmith and Lakoff's papers,

30 Epenthesis 1 and Voicing Assimilation apply simultaneously at the M-W component.

the relationship of well-formedness conditions holding at each level is not given due consideration. Goldsmith (1993b: 24), for instance, remarks that "Levels, indeed, may be quite autonomous and independent of each other." Is this really the case with respect to phonotactics at each level? For example, what is the relationship between W-level phonotactics and P-level phonotactics? It seems natural to assume that W-level phonotactics should be satisfied at level P as well; however, unless we stated something along these lines, the theory could allow a possibility that once we satisfy W-level phonotactics, cross-level (W, P) constructions may apply to violate them.

This issue is relevant to the status of phonotactics (8) and (8)', repeated below:

(8) $*[[+\text{voi}]\ [+\text{voi}]]_{\text{morpheme}}$
(8)' $*[[+\text{voi}]\ [+\text{voi}]]_{\text{word}}$

If levels are really independent of each other, as Goldsmith seems to suggest, then we have to state the phonotactic twice, first at level M and second at level W. This duplication can in principle be avoided if we assume that M-level phonotactics are observed at level W unless otherwise specified.

The second problem of the proposed analysis concerns the formulation of Gemination, which is repeated below:

(25) Gemination
M: C_i
   |
W: C_j C_j

The past tense formation of /b/-final verb stems was analyzed as follows in Section 4:

(19c)' M: tob + ta
(Voicing Assimilation, Gemination)
W: todda
(Voiced Geminate Nasalization)
toNda
(Nasal Assimilation)
P: tonda

The reason the environment statement is made at level W for Gemination is that otherwise the past tense formation of /b/-final verb stems would not be properly explained. Notice that if the environment condition were stated at level M as in (25)', simultaneous application of Voicing Assimilation and Gemination could change *tob + ta* to *totda*, instead of *todda*, at level W:

(25)' M: C_i C_j
|
W: C_j

If we were to maintain (25)', we might consider the following two as alternative analyses:

(48) M: tob + ta
(Voicing Assimilation)
W: tobda
(Gemination)
P: todda
(Voiced Geminate Nasalization)
toNda
(Nasal Assimilation)
tonda

(49) M: tob + ta
(Voicing Assimilation)
W: tobda
(Gemination)
todda
(Voiced Geminate Nasalization)
toNda
(Nasal Assimilation)
P: tonda

(48) is infeasible simply because the form is not properly syllabified at level W. (49) appears more promising than (48); however, the idea of Gemination as an intralevel construction is not appealing. As we saw in Section 2.1, an intralevel construction is basically a repair strategy that fixes up ill-formedness caused by a cross-level construction. In (49), it is not Voicing Assimilation that causes ill-formedness; rather, it is the M-level consonant cluster itself that needs to be fixed up by Gemination. For this reason, I would reject (49), too. This brings us back to the original analysis shown in (19c)', together with Gemination (25), where the environment condition is stated at level W. Cognitive phonology theoretically allows one to make environment statements at either level of a cross-level construction (cf. Lakoff 1993: 120). If, however, one is allowed to decide where to make environment statements just to "squeeze" all the relevant constructions within the three levels, this is a serious drawback of the theory. It is necessary to consider if there are any constraints on where (i.e. at which level) to make environment statements.

The third problem concerns how to incorporate Zec's (1993) Moraic Prominence constraint into our analysis. The relevant question here is, is it a mere coincidence that Gemination and Epenthesis preserve the number of moras? Would it be possible to have a construction that would change *tor* + *ta* to *tota*? How can we constrain constructions in terms of the number of moras? I suspect that a possible solution might lie in the idea of "Default Identity", which is defined as "Identity across all levels is maximally harmonic except where harmony is increased by the fitting of constructions" (Lakoff 1993: 141). Following this idea, we might say that the identity of the number of moras across all levels is maximally harmonic.

The fourth problem concerns the measurement of harmony. In harmonic application, rules apply in such a way as to shift the representations toward the state that maximally satisfies the phonotactics of the relevant level; however, it is not clear how we should assess one representation as more "satisfactory" than the other.[31] Let us consider (19c)' again as an example. Voiced Geminate Nasalization is an intralevel construction, hence applying harmonically. The representation *todda*, however, could have shifted to *totta*, since the latter perfectly satisfies the phonotactics at level W. How should we decide that *DD → ND is more harmonic than *DD → TT? Prince and Smolensky (1993: 123) criticize Goldsmith's analysis of Lardil as relying on "pre-theoretic notions of harmony" and I agree with them that cognitive phonology needs to explicitly define how to measure degrees of harmony (cf. Sommerstein 1974: 75–76).

6. Conclusion

This paper has been an attempt to explore possibilities and problems of cognitive phonology. We saw in Section 2 that cognitive phonology is a cognitively viable approach based on connectionist foundations, offering an alternative view to the assumptions of classical generative phonology. Sections 3 and 4 applied cognitive phonology to the past tense formation of Japanese verbs, showing that, despite some problems, the proposed analysis can deal with the data without imposing extrinsic rule ordering and that this construction-based approach makes it possible to account for both purely phonological alternations and morphologized alternations.

Despite some vagueness and problems of the framework, cognitive phonology offers an interesting perspective on phonology. I hope the present paper has at least shown that its possibilities are worth investigating.

REFERENCES

Anderson, Stephen R. and Brent de Chene (1979) "Compensatory lengthening." *Language* 55: 505–535.

Ashworth, David (1976–77) "Historical perspective on voicing and the gerund in Japanese." *Papers in Japanese Linguistics* 5: 27–39.

Bybee, Joan (2001) *Phonology and Language Use*. Cambridge: Cambridge University Press.

Davis, Stuart and Natsuko Tsujimura (1991) "An autosegmental account of Japanese verbal conjugation." *Journal of Japanese Linguistics* 13: 117–144.

Goldsmith, John (1989) "Licensing, inalterability, and harmonic rule application." *CLS* 25: 145–156.

____. (1990) *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.

____. (1991) "Phonology as an intelligent system." In Donna J. Napoli and Judy Kegl eds., *Bridges Between Psychology and Linguistics: A Swarthmore Festschrift for Lila Gleitman*, 247–267, Hillsdale, N. J.: Lawrence

31 Goldsmith (1993b: 28) states that "each level also contains complexity measures, in such a fashion that certain representations on a given level may be said to be more complex than others on the same level", but he does not specifically mention how to define or measure "complexity".

Erlbaum.
____. (1993a) "Introduction." In J. Goldsmith ed. (1993), 1–20.
____. (1993b) "Harmonic phonology." In J. Goldsmith ed. (1993), 21–60.
____. ed. (1993) *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago and London: The University of Chicago Press.
Hayes, Bruce (1986) "Inalterability in CV phonology." *Language* 62: 321–351.
Itô, Junko (1986 [1988]) *Syllable Theory in Prosodic Phonology*. New York/London: Garland Publishing, Inc.
Itô, Junko and Ralf-Armin Mester (1986) "The phonology of voicing in Japanese: theoretical consequences for morphological accessibility." *Linguistic Inquiry* 17: 46–73
____. (1993) "Japanese phonology: constraint domains and structure preservation." LRC-93-06, The Linguistics Research Center, University of California, Santa Cruz.
Kenstowicz, Michael and Charles Kisseberth (1979) *Generative Phonology*. San Diego: Academic Press, Inc.
Kiparsky, Paul (1973) "Elsewhere in phonology." In Stephen Anderson and Paul Kiparsky eds., *A Festschrift for Morris Halle*, 93–106. New York: Holt Rinehart and Winston.
Kumashiro, Fumiko (2000) *Phonotactic Interactions: A Non-Reductionist Approach to Phonology*. Doctoral dissertation, University of California at San Diego.
Kuroda, S.-Y. (1965 [1979]) *Generative Grammatical Studies in the Japanese Language*. New York/London: Garland. Publishing, Inc.
____. (1966) "Soku-on oyobi hatu-on ni tuite" (On the mora obstruent and mora nasal). *Gengo Kenkyu* 50: 85–99.
Lakoff, George (1988) "A suggestion for a linguist with connectionist foundations." In D. Touretzky et al. eds., *Proceedings of the 1988 Connectionist Models Summer School*, 301–314, Morgan Kaufmann.
____. (1993) "Cognitive phonology." In J. Goldsmith ed. (1993), 117–145.
Langacker, Ronald W. (1987a) *Foundations of Cognitive Grammar*, Vol. 1. Stanford: Stanford University Press.
____. (1987b) "The cognitive perspective." *Center for Research in Language Newsletter* 1.3: 4–14.
____. (1988) "A usage-based model." In Brygida Rudzka-Ostyn ed. *Topics in Cognitive Linguistics*, 127–161. Amsterdam/Philadelphia: John Benjamins.
McCawley, James (1968) *The Phonological Component of a Grammar of Japanese*. The Hague/Paris: Mouton.
Nathan, Geoffrey S. (2008) *Phonology: A Cognitive Grammar Introduction*. Amsterdam/Philadelphia: John Benjamins.
Poser, William (1986) "Japanese evidence bearing on the compensatory lengthening controversy." In L. Wetzels and E. Sezer eds., *Studies in Compensatory Lengthening*, 167–186. Dordrecht/Riverton: Foris.
Prince, Alan and Paul Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report ♯2 of the Rutgers Center for Cognitive Science.
Rumelhart, David E, James L. McClelland eds. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*. Cambridge: The MIT Press.
Schane, Sanford A. (1971) "The phonemes revisited." *Language* 47: 503–521.
Smolensky, Paul (1986) "Information processing in dynamical systems: foundations of harmony theory". In Rumelhart and McClelland eds. (1986), 194–281.
Sommerstein, Alan H. (1974) "On phonotactically motivated rules." *Journal of Linguistics* 10: 71–94.
Vance, Timothy (1987) *An Introduction to Japanese Phonology*. Albany: State University of New York Press.
Wiltshire, Caroline R. (1992) *Syllabification and Rule Application in Harmonic Phonology*. Docoral dissertation, University of Chicago.
Yoshiba, Hiroshi (1981) "The mora constraint in Japanese phonology." *Linguistic Analysis* 7: 241–262.
Zec, Draga (1993) "Patterns of gemination and consonant loss: Pali, Japanese, and cross- linguistic." Rutgers Optimality Workshop handout.