論文

The Competitive Relationship between Japanese Accent and Intonation

Toshiyuki Sadanobu

要旨：
　本稿は、現実のコミュニケーションを考慮に入れることによって、現代日本語（共通語）の話しことばに対する我々の理解を、特に語彙的なアクセントとフレーザルイントネーションの関係に関して深めようとするものである。

　伝統的には、共通語における語彙的アクセントとフレーザルなイントネーションは、基本的に非競合的と考えられてきた。だが、この考えでは、連なり型や重なり型は説明できるが、排他型は説明できない。排他型はこれまでほとんど取り上げられていなかったが、よく注目してみると、いくつかの方言と同様、これまで伝統的に考えられていたよりも、頻繁に幅広く観察できる。そこで本稿では、語彙アクセントとイントネーションの関係について、競合説を提案する。結論は以下4点である。

　第1点：日本語（共通語）の語彙的アクセントとフレーザルなイントネーションの関係をとらえるには、伝統的な非競合説よりも競合説の方が有効である。アクセントとイントネーションは潜在的には、ピッチへの反映をめぐって潜在的に競合している。つまり基本は排他型である。

　第2点：非競合説よりも競合説の方がよいという本稿の主張は、以下2つの観察結果により支持される。観察1：これまで考えられていたよりもずっと頻繁に幅広く生じることが判明した排他型を説明できるのは競合説だけである。観察2：連なり型や重なり型の場合には話し手の強い気持ちがなく、このことを説明するには競合説に立ち、「強さのアイコニシティ」の考え（話し手の気持ちが強いほど、その気持ちに対応する韻律が強い力

を得て、排他型を生み出しやすくなるという考え）を取り入れる必要があ
る。

　第 3 点：強さのアイコニシティは、少なくとも 2 点の根拠を持つ。それ
は、疑似的な最小対立を用いた知覚実験と、アクセントがイントネーショ
ンを抑え込む現象の存在である。

　第 4 点：これまで非競合説が基本と考えられていたのはおそらく、実験
室的環境では被験者はふつう強い気持ちを示さないせいだろう。

**キーワード：日本語、アクセント、イントネーション、競合説、強さのア
イコニシティ**

Abstract：

　This paper aims to deepen our understanding of grammar of spoken Standard
Japanese with special reference to the relationship between lexical accent and
phrasal intonation by taking actual communication into consideration.

　Traditionally, lexical accent and phrasal intonation in Standard Japanese have
been thought as non-competitive. By this view, there are no competition for
realization between lexical accent and phrasal intonation. Such a view of non-
competition can explain copulative or cumulative forms, whereas it cannot
explain conflictive form. Although research on Standard Japanese have been
reluctant to admit conflictive forms, close investigation reveals that conflictive
form is much more often and widespread than has traditionally been thought
as well as in some of the other dialects. This paper suggests a new view which
can explain all of these three pitch forms. Conclusions are as follow.

　First, the relationship between lexical accent and phrasal intonation in
Standard Japanese can be better understood by competitive view rather than by
traditional non-competitive view. Lexical accent and intonation are potentially
competing for generating their own pitch form (i.e. conflictive form).

　Second, the superiority of competitive view over non-competitive view
is supported by two observations: (i) It is only competitive view that can

explain conflictive form, which is much more often and widespread than has traditionally been thought; (ii) By adopting Iconicity of Strength (i.e. The stronger the speaker's attitude is, the more likely the corresponding prosody has strong power to generate a conflictive form), competitive view rightly expects lack of strong attitudes in copulative and cumulative forms.

Third, Iconicity of Strength can be justified at least in two ways: one is perceptual test, and the other is the possibility of upset victory of accent over intonation.

And fourth, non-conflictive forms have traditionally been thought as basic, probably because it is rare for informants to show their strong attitudes and prominent characters in the environment of laboratory.

1. Introduction

In this paper I aim to deepen our understanding of grammar of spoken Standard Japanese with special reference to the relationship between lexical accent and phrasal intonation by taking actual communication into consideration. Traditionally, lexical accent and phrasal intonation in Standard Japanese have been thought of as non-competitive. This view, however, has many more problems than has traditionally been thought. This paper suggests a new competitive view which can explain all three pitch forms
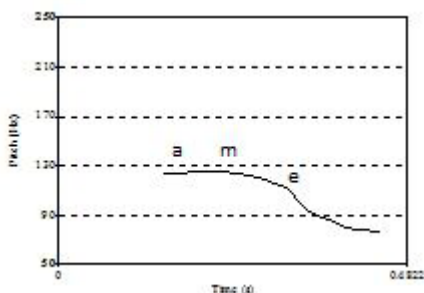
Previous research based on the non-competitive view is outlined in Section 2, and its defects are examined in Section 3. In Section 4 I shall show that the competitive view is more preferable than the non-competitive one, and in Section 5 I shall raise counter-arguments against the competitive view and refute them. Finally I conclude my arguments in Section 6.

2. Previous research

Traditionally, lexical accent and phrasal intonation in Standard Japanese have been basically thought of as non-competitive. By this view, there is no competition for realization between lexical accent and phrasal intonation. For example, Amanuma *et al*. (1978) says that intonation cannot break patterns of lexical accents in Japanese language. Another example is Abe (1998) which says "(Japanese lexical accent is) susceptible but usually not subservient to intonation... (It) resists being perturbed by intonation." See the introductory part of Sadanobu (2005a) for more details of the traditional view.

According to the traditional view ("non-competitive view," henceforth), lexical accent and phrasal intonation jointly generate either one of two pitch forms; "copulative" or "cumulative."
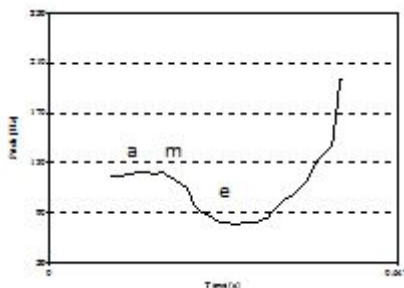
By "copulative," I mean here a pitch form where a given accent is accompanied by intonation in this order. Let us take a word for example. The lexical accent of the word *ame* (rain) is *a* (High) – *me* (Low), which can be ascertained when we pronounce this word isolated from other parts of utterance (Sound 1, Figure 1) (Note 1). (I use praat (praat 5207, http://www. fon.hum.uva.nl/praat/download_win.html) for extraction and analysis of sound. The same hereinafter unless otherwise noted.)

Sound 1: Isolated pronunciation of the word *ame* (rain)

Figure 1: Pitch of Sound 1

And the pitch of the interrogative sentence *ame*? (Rain?) consists of two parts. The former part reflects the lexical pattern (High-Low) of the word *ame*, and the latter part reflects the phrasal pattern (rising) of the interrogative intonation (Sound 2, Figure 2).
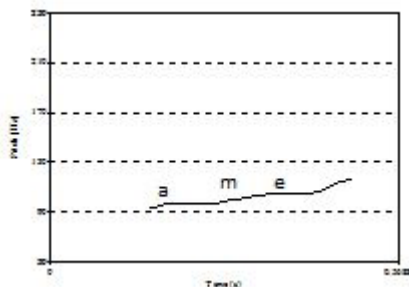

Sound 2: Interrogative sentence *ame*?　　　　**Figure 2: Pitch of Sound 2**

This is why we can say that the relationship between lexical accent and phrasal intonation of *ame*? (Rain?) is copulative
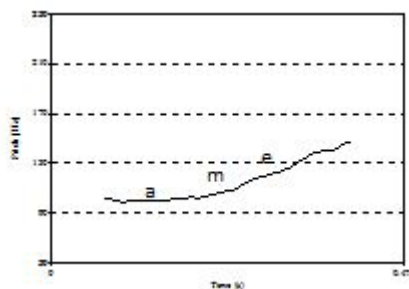
When the directions of accent and intonation are the same, rising for example, a cumulative form often appears. By "cumulative" I mean a pitch form which is affected by intonation (rising), and becomes more rising. For example, the lexical accent of the word *ame* (candy) is *a* (Low) – *me* (High) (Sound 3, Figure 3).

Sound 3: Isolated pronunciation of the word *ame* (candy)

Figure 3: Pitch of Sound 1

And its rising pattern becomes more remarkable in the interrogative sentence *ame*? (Candy?) (Sound 4, Figure 4).

Sound 4: Interrogative *ame*? (Candy?)

Figure 4: Pitch of Sound 4

Therefore we can say that the relationship between the lexical accent and phrasal intonation of *ame*? (Candy?) is cumulative.

Copulative and cumulative forms introduced above can be explained by the non-competitive view as a matter of course. On the other hand, there remains a pitch pattern which the non-competitive view cannot explain by nature. This is what we call the conflictive form (Table 1).

Table 1: The non-competitive view can account for copulative and cumulative forms but not conflictive forms.

	copulative	cumulative	conflictive
Non-competitive view	+	+	−

By a conflictive form, I mean a pitch pattern which follows either lexical accent or phrasal intonation. In most cases of the conflictive form, accent is affected by intonation to such a degree that it has completely lost its original shape. For example, the lexical accent of the word nani (what) is na (High) – ni (Low) (Sound 5, Figure 5).

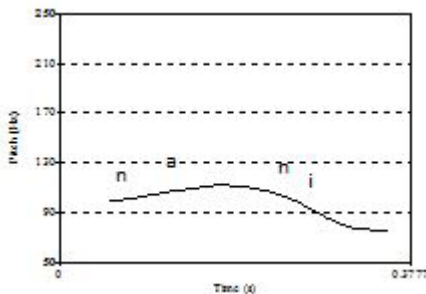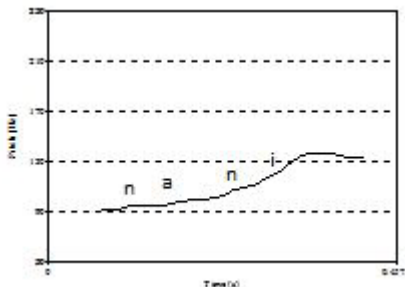**Sound 5: Isolated pronunciation of the word *nani* (what)**

Figure 5: Pitch of Sound 1

However, this word can be uttered in rising pitch with a strong attitude of asking the question. (Sound 6, Figure 6).

Sound 6: Strong question
***nani*? (What?)**

Figure 6: Pitch of Sound 6

In this case we can say that lexical accent and phrasal intonation are conflictive. Although the pitch pattern slightly descends at the last part in Fig. 6, it does not change our judgment of this pitch pattern as conflictive. This is because such a subtle fall of pitch at the last part of the utterance can be seen in natural data as well (e.g. *oishii naa* (Sound 20, Figure 10) shown in Section 4.1 below) which we feel as conflictive.

　Research on Standard Japanese has been reluctant to admit conflictive forms. For example, Amanuma et al. (1978) pointed out the following three patterns "V-*tai*," "V/A/N-*desho*?," and "V-*tekudasai*" as conflictive, but there remain many unclear points. Firstly, the pattern V-*tai* can be pronounced not only in the conflictive form like "*tabetai*" but also in the copulative form like "*tabetai*." So what is the difference between "*tabetai*" and "*tabetai*"? Secondly, what is the difference between conflictive "*taberu deshoo*" and copulative "*taberu deshoo*"? And thirdly, when and why do we pronounce this pattern in the conflictive form, like "*tabete kudasai*"? And why is it more natural to pronounce this pattern in the conflictive form with the sentence-final particle "*yo*," as in "*tabete kudasai yoo*"? Such questions all remain unanswered. The situation is similar for other studies.

　It is not right that there has been no description that admits the competitive relationship between lexical accent and phrasal intonation (Sadanobu 2005a).

Kawakami (1963: 33, 37) notes "Accentual nuclear disappears because of rising tone," and "when the speaker utters very lightly with rising intonation (second type), s/he must delete the accentual fall." Akinaga (1966: 58–59) admits that tones of respectful reading and of recitation and the tone of child actor/actress's speaking "ignores" and "destroys" accentual patterns, although it is not clear whether these tones belong to intonation. Also Moriyama (1989: 173–174) points out that intonation can delete accentual fall with limitations of word length and accentual patterns. Isamu Abe seems to have admitted the conflictive form for a long time, and this paper borrows the English terms "copulative," "cumulative," and "conflictive" from Abe (1998: 362). Kori (1997: 172–184) states that the intonation of a focused word can weaken the accents of following words. Though distinct from the above research, this is also regarded as one of the descriptions that admits that the reflection of lexical accent onto pitch can be disturbed by intonation. However, such descriptions that admit the conflictive form more or less are not so large in number. Are conflictive forms so rare?
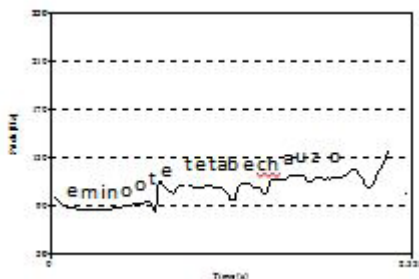
3. Some observations

Turning to dialectal research, we can find conflictive forms much more commonly. According to Fujiwara (1997), for example, Southern Aso dialect has a conflictive form of a high flat type in which the speaker can pronounce a sentence "*Aaga shigoto wa daa mo sen*" (Nobody undertakes that work) with no lexical accent and a high flat pitch. Also Kyoto dialect has a conflictive form of a meandering type in which a sentence such as "*Sonna koto shitara aka hen yanai no*" (You're a fool, to do such a thing!) can be pronounced with a high pitch (for *sonna*, *shi*, *a*, and *hen*) and a low pitch (for *koto*, *tara*, *ka*, and *yanaino*).

Such a difference between research of dialects on one hand and of Standard Japanese on the other hand might be attributed to the qualitative difference of data observed. What is commonly observed in dialectal research and often

ignored in (laboratory-like) Standard Japanese research are utterances made with a strong attitude. And by focusing on these, we can find conflictive forms also in Standard Japanese. Below I shall show three examples.

The first example of a conflictive form in Standard Japanese is a gradual rising type (Sound 7, Figure 7) (Note 2).

Sound 7: *Emi no otete tabechau zo waan!* **(It will eat your hand, Emi. Wham!)**　　**Figure 7: Pitch of Sound 7**

This sound is taken from the so-called "Emichan data," a mother-daughter conversation database recorded in the latter half of the 1970s (Sugito 2005). (Vol.14, Side B, around 49: 32) (The reason I use this database in spite of its oldness is that it is actually not old at all, at least for its prosodic aspect.) And here is an utterance of a mother who is trying to frighten her daughter, Emi, by joking. At the last part she even cries an onomatopoeia *Waan* (Wham), a sound mimicking something (maybe a monster judging from the context, or a big fish judging from the conversation of Vol.11, Side B, around 15: 35) trying to eat Emi's hand. Apart from this cry, the pitch of this utterance gradually gets higher and higher from around 90Hz to 110Hz in about 2 seconds. We cannot admit any reflection of lexical accent of words such as *Emi* (High-Low) and *tabechau* (High-Low-Low-Low) there.

The second example of a conflictive form is a gradual falling type (Video 1, Sound 8, Figure 8).

Video 1: KinKi Kids

Sound 8: *Hidoi yo* (How cruel!) from Video 1　　　**Figure 8: Pitch of Sound 8**

This is taken from a sound database of everyday conversation collected by my research group in the mid 2000s. Here two young female speakers are talking about a famous idol group *KinKi Kids*. According to my auditory impression, the utterance "*Hidoi yo*" (How cruel!) gradually descends (around from 380Hz to 330Hz in Figure 8) over about 1.5 seconds. Here the lexical accent of the word *hidoi* (cruel, Low-High-Low) is not reflected to pitch. (Although the falling pitch shape of "*Hidoi yo*" in Figure 8 is not so clear because of overlapping with the other speaker's voice saying "*KinKi, KinKi*," I think the idea that nothing can be said unless you have a clear reflection on the pitch figure is unrealistic.)

The third and last example of a conflictive form is the "descent from initial high pitch" type (Video 2, Sound 9, Figure 9).

Video 2: After telephone conversation

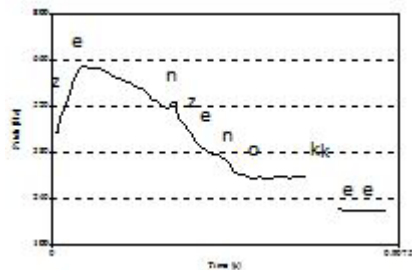**Sound 9:** *Zenzen okke* **(No problem) from Video 2**

Figure 9: Pitch of Sound 9

This data is taken from an audio-visual database of everyday conversation collected by my research group in the late 2000s. The video was captured by a camera with a 360 degree visual field, and its upper and lower halves cover 180 degree angles respectively. (The leftmost end of the upper part is connected with the rightmost end of the lower part, and the rightmost end of the upper part with the leftmost end of the lower part.) Here we can see the interior space of a house. On the upper field a woman stands with her face away from

us, talking with someone by telephone. Another woman on the lower field is sitting with her profile directed toward us. After the first woman finishes the telephone conversation in a polite tone, the second woman asks her "*Daijobu deshita?*" (Was it ok?), and the first woman answers "*Zenzen okke*" (No problem!) with the initial part, ze, pronounced as high as 640Hz and the remaining part pronounced with a falling pitch. The lexical accent of the word *zenzen* (at all, Low-High-High-High) is not reflected to pitch.

In all three cases, there is no reflection of lexical accent in pitch. Thus the conflictive form is much more common and widespread than has traditionally been thought.

4. Competitive view and "iconicity of strength"

Based on the observations above, I shall suggest a competitive view for the relationship between lexical accent and phrasal intonation in Standard Japanese. According to this view, the relationship between lexical accent and phrasal intonation is basically conflictive. That is to say, they are potentially competing with each other to generate their own pitch form (i.e. conflictive form).

It goes without saying that the competitive view can explain conflictive forms. So how can copulative and cumulative forms be explained?

I shall show that the competitive view can explain these non-conflictive forms as well as conflictive forms by using the "iconicity of strength" hypothesis. This hypothesis can also characterize conflictive and non-conflictive forms respectively.

The "iconicity of strength" hypothesis (hereafter IS) expresses the following idea: The stronger the speaker's attitude is, the more likely the speaker is to concentrate on producing the corresponding prosody. As a result, the prosody has enough power to generate conflictive forms.
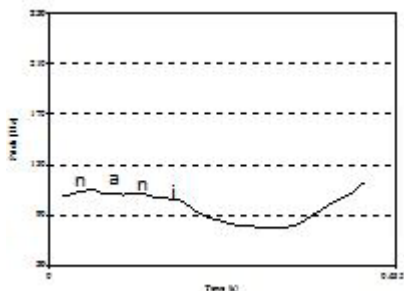
If we adopt IS, copulative and cumulative forms are characterized as forms

lacking the speaker's strong attitude, distinct from the conflictive form, which has such strong attitudes. Copulative and cumulative forms can be understood as "drawn ties" between accent and intonation because of this lack of strong attitudes.

Below I shall show the validity of IS in two ways. One is a perceptual test and the other is the fact that lexical accent is not always the loser and can defeat phrasal intonation.
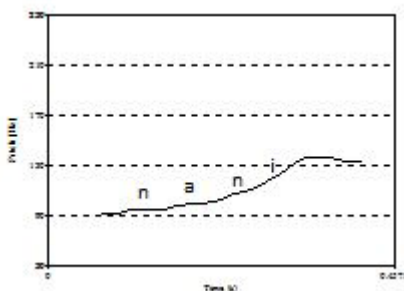
4.1 Perceptual test

The first way of showing the validity of IS is by conducting a perceptual test using quasi-minimal pairs. By quasi-minimal pair, I mean here a pair such as the following two utterances of "*nani?*" (What?), one of which is copulative (Sound 10, Figure 10) and the other of which is conflictive (Sound 11, same as Sound 6, Figure 11).

Sound 10: *nani?* (What?) in copulative form

Figure 10: Pitch of Sound 10

Figure 11: Pitch of Sound 11

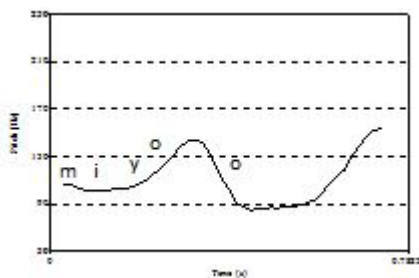**Sound 11: *nani?* (What?) in conflictive form (= Sound 6)**

If these two utterances differed from each other only in pitch, we could call them a minimal pair. Actually they are just a quasi-minimal pair, since they differ not only in pitch but also in power, duration, voice quality etc. (There is a big difference of duration at least.) Although the ideal pair for a perceptual test should be a minimal pair, this is virtually unobtainable as far as we focus on human raw sound and refrain from using signal processing technology, like this paper. This is the reason for our compromise of using quasi-minimal pairs here.

The subjects of the perceptual test were 30 university students in the Kansai area and they were not paid. They listened to each utterance of quasi-minimal pairs twice and were asked to detect the speaker's attitude. Their answers were made by selecting candidates of attitude already given to them. It was announced to the subjects before the test that they could select more than one candidate if they felt like doing so. And it was also announced to them that they could select no candidate if they didn't find any suitable one corresponding with the utterance.

The result of the test was as follows: In the case of copulative nani? (What?), as many as 29 subjects selected the candidate "asking" and three subjects selected "weak surprise." There were no subjects who selected "strong
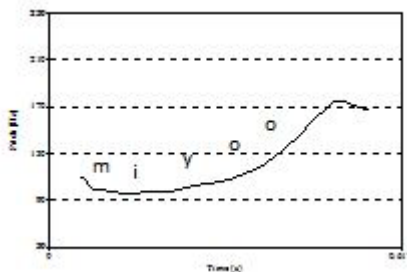
surprise." On the contrary, in the case of the conflictive form, 28 subjects selected "strong surprise," 17 subjects "weak surprise," and no subjects "asking." If we accept the common sense that of these three attitudes "strong surprise" is the strongest, "asking" is the weakest, and "weak surprise" is intermediate between them, we can understand this attitudinal difference between copulative and conflictive forms with IS in the following way: Rising intonation connected with strong attitudes has a power strong enough to generate its own pitch form. (Note 3)

A similar result was gained in a quasi-minimal pair with verb stem + yoo ending. For example, "miyoo?" (Watch?) in the copulative form (Sound 12, Figure 12) is likely to be interpreted as reasking such as "Did you say 'watch'?" (by 24 subjects) and unlikely to be interpreted as inductive such as "Shall we watch together?" (no subjects). Quite contrary to this, miyoo? (Watch?) in the conflictive form (Sound 13, Figure 13) is unlikely to be interpreted as reasking (except 1 subject) and likely to be interpreted as inductive (by 27 subjects).

Sound 12: *miyoo?* (Watch?)
in copulative form
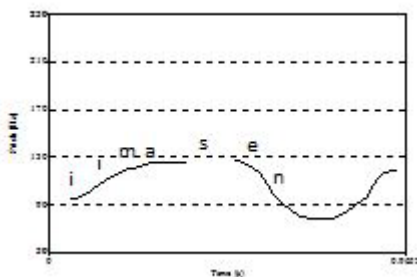
Figure 12: Pitch of Sound 12

Sound 13: *miyoo?* (Watch?)　　　**Figure 13: Pitch of Sound 13**
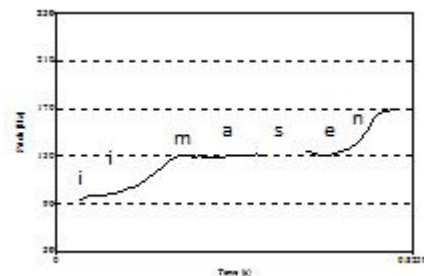**in conflictive form**

Such an attitudinal difference between copulative and conflictive forms of *miyoo?* (Watch?) can be understand by IS as follows: Compared with the rising intonation of reasking utterances, that of inductive utterances has stronger power for generating its own pitch form because the attitude of induction is deontic and stronger than reasking.

We can see the same trend also in the case of quasi-minimal pairs with verb stem + *masen* endings, although the connection between pitch form and attitude is somewhat overlapping. For example, "*iimasen?*" (Do not say?) in the copulative form (Sound 14, Figure 14) tends to be interpreted as reasking ("Did you say 'I don't say'?", 17 subjects) and confirmation ("Is it right that you don't say?", 24 subjects), rather than as strong confirmation ("You say, don't you?", 5 subjects) and inducement ("Let's say together, 5 subjects). By contrast, iimasen? (Do not say?) in the conflictive form (Sound 15, Figure 15) tends to be interpreted as inducement (28 subjects) and strong confirmation (20 subjects) rather than as confirmation (8 subjects) and reasking (1 subject).

Sound 14: *iimasen?* (Do not say?) in copulative form

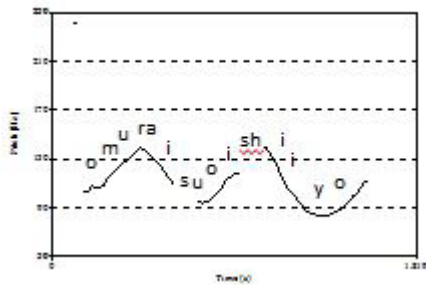**Figure 14: Pitch of Sound 14**

Sound 15: *iimasen?* (Do not say?) in conflictive form

Figure 15: Pitch of Sound 15

This difference between copulative and conflictive forms of *iimasen?* (Do not say?) is quite naturally understood using IS once we admit the hierarchy of attitudinal strength among reasking, confirmation, strong confirmation, and inducement as the attitude gets stronger.
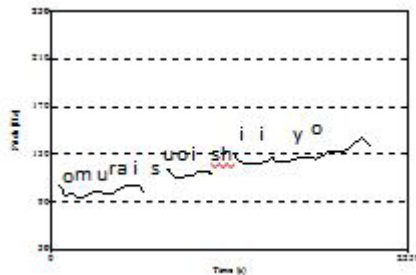
This is true also in the case of sentences with sentence-final particles. Let us take a sentence "*omuraisu oishii yo*" (Rice omelets taste good.) with the sentence-final particle *yo* as an example. When this sentence is pronounced in the copulative form (Sound 16, Figure 16), it is apt to be interpreted widely not only as an utterance of relatively weak attitudes such as reasking "Did

you say 'omuraisu oishii yo'?" (20 subjects) and asking "Do you know rice omelets taste good?" (20 subjects) but also as an utterance of relatively strong attitudes such as inducement "Let's have rice omelets together." (15 subjects) and jokingly threatening "You will regret it if you have a rice omelet." (13 subjects). By contrast when it is pronounced in the conflictive form (Sound 17, Figure 17), it is interpreted only as utterance of inducement (28 subjects) and jokingly threatening (28 subjects).

Sound 16: *Omuraisu oishii yo* (Rice omelets taste good.) in copulative form
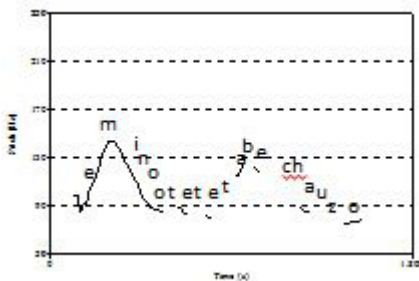
Figure 16: Pitch of Sound 16

Sound 17: *Omuraisu oishii yo* (Rice omelets taste good.) in conflictive form
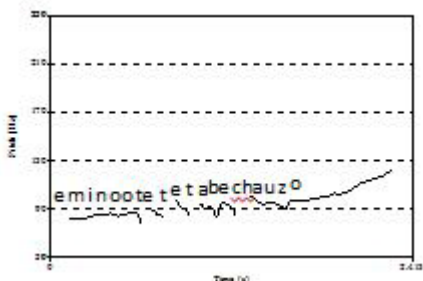
Figure 17: Pitch of Sound 17

It is noteworthy here that conflictive forms like those in Sound 17 become unnatural when we drop the sentence-final particle *yo*. That is to say, the sentence-final particle *yo* makes the conflictive form more natural. Although the semantics of *yo* are not easy to capture, we can safely say that sentence-final particles such as *yo* contribute to the expression of the speaker's strong attitudes as a result. Thus we can understand why sentence-final particles facilitate the occurrence of conflictive pitch by using IS.

As well as *yo*, the sentence-final particle *zo* contributes an expression of the speaker's strong attitudes and therefore enables conflictive pitch form. Let us take the sentence "*emi no otete tabechau zo*" (It will eat your hand, Emi.) for example. If pronounced in the copulative form (Sound 18, Figure 18), the subjects tend to interpret this sentence as an utterance of confirmation "Do you understand that it will eat your hand, Emi?" (12 subjects) and of threatening "I hereby threaten that it will eat your hand, Emi." (18 subjects). And if this sentence is pronounced in the conflictive form (Sound 19, Figure 19), it is interpreted only as an utterance of threatening (27 subjects) and no subjects interpreted this sentence as an utterance of confirmation.

Sound 18: *emi no otete tabechau zo* **(It will eat your hand, Emi.) in copulative form**

Figure 18: Pitch of Sound 18

Sound 19: *emi no otete tabechau zo* (It will eat your hand, Emi.) in conflictive form

Figure 19: Pitch of Sound 19

The "Emichan data" has not only *zo* sentences but also *na* sentences uttered in a conflictive pitch form. Let us see a scene where Emi and her mother are reading a picture book together. There Emi's mother pretends to eat a rice ball in the book and exclaims the sentence "*oishii na*" (How delicious!) in the conflictive form (Vol. 13, Side B, 30: 20, Sound 20, Figure 20).
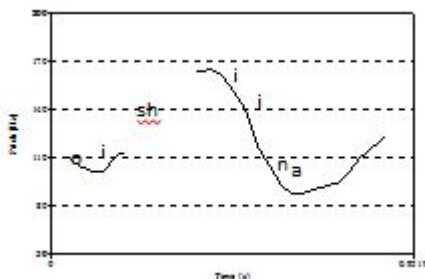

Sound 20: *oishii na* (How delicious!, natural utterance)
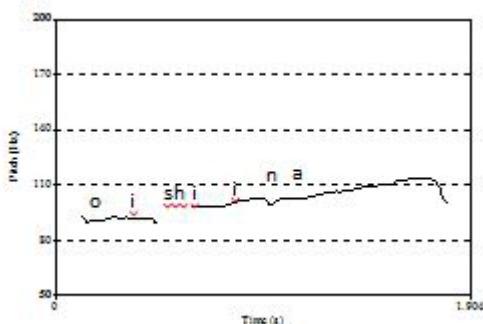
Figure 20: Pitch of Sound 20

After preparing a quasi-minimal pair of the copulative *oishii na* (Sound 21, Figure 21) and corresponding conflictive *oishii na* (Sound 22, Figure 22), a perceptual test was conducted. The result was that *oishii na* in the

copulative form tends to be interpreted as an utterance of confirmation (10 subjects) and that of reasking (5 subjects) rather than that of exclamation (0 subjects), whereas *oishii na* in the conflictive form is likely to be interpreted as exclamatory utterance rather as confirmation (3 subjects) and reasking (0 subject).

Sound 21: *oishii na* (How delicious!) in copulative form

Figure 21: Pitch of Sound 21

Sound 22: *oishii na* (How delicious!) in conflictive form
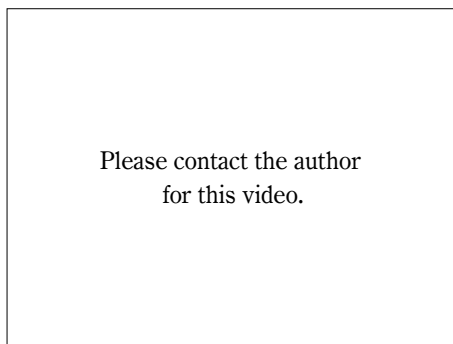
Figure 22: Pitch of Sound 22

4.2 Lexical accent can defeat phrasal intonation

IS has other evidence for its validity: The competition between lexical

accent and phrasal intonation for pitch reflection is not always one-sided. It is rare but possible for lexical accent to defeat phrasal intonation. And when and why it occurs can be understood only by using IS.

Let us take for example a scene from a TV broadcast of a motor race (*F1 Grand Prix in Monaco*, 2003 May 31, Kansai TV). In this scene a racer named Montoya enters his car into his team's (Williams BMW) pit to change tires and refuel (Video 3).

Please contact the author
for this video.

Video 3: TV Broadcasting of Montoya's pit stop.

At the last part of this scene, the broadcaster (Mr. Tsuneo Shiobara) makes the utterance "*Saa, hachibyou ni toyuu seishi jikan no ato, supiido seigen no moukerareta kono semai pitto roodo o ikimasu Pan Paburo Montooya no mashiin!*" (Now, after an 8.2 second stop, Juan Pablo Montoya's car goes along this narrow, speed-limited pit road!) with high flat pitch pattern for the most part (Sound 23, Figure 23. For the analysis of Sound 23, which has much engine noise, I owe Zhu Chunyue who made Figure 11 by using SUGI Speech Analyzer, software developed by Animo and supervised by Miyoko Sugito. http://www.animo.co.jp/analyze/sugi/). (The racer referred to by the broadcaster as *Pan Paburo Montooya* should be rightly pronounced as Hoan Paburo Montooya (*Juan Pablo Montoya*), but here I transcribed his name straight from my auditory impression of the broadcasting.)

71

Sound 23: Utterance *saa, hachibyou ni toiu seishi jikan no ato, supiido seigen no moukerareta kono semai pitto roodo o ikimasu Pan Paburo Montooya no mashiin!* **(Now, after an 8.2 second stop, Juan Pablo Montoya's car goes along this narrow, speed-limited pit road!)**

Figure 23: Pitch of Sound 23

Here we can see the conflictive pitch form. A high flat intonation connected with a strong "crying" excited attitude expels lexical accents and governs almost all parts of the utterance. However, for the part *semai* (narrow), there is a clear up and down of pitch, reflecting its lexical accent (Low-High-Low) (Figure 23, red circle). This means that lexical accent of *semai* withstands and suppresses the high flat intonation. Why is such a reversal upset brought about? By admitting IS, we can understand this in the following way: In the preceding context the broadcaster conversed with the guest (a former F1 racer, Ukyo Katayama) and the reporter (Masahiko Kondo) about how narrow the pit road is, and so the word *semai* was connected with his strong (emphatic) attitude, which gave strong enough power to the lexical accent to suppress the high flat intonation.

In Video 3 above, we can ascertain how often they talked about the narrowness of the pit road before the time point of uttering Sound 23. And at the end of Video 3, (immediately after Sound 23), we can hear that the reporter

Masahiko Kondo cries "*deguchi mo semai!*" (The exit is also narrow!)in the conflictive form, which implies that the narrowness of the pit road remained as a potential topic while they broadcasted Montoya's pit-stop.

What I discussed regarding the non-competitive view and competitive view above is summarized in the following two points: Firstly, unlike the non-competitive view, the competitive view can explain the conflictive form, which is much more common and widespread (including the victory of lexical accent over phrasal intonation) than has traditionally been thought. Secondly, by adopting Iconicity of Strength, the competitive view rightly expects the lack of strong attitudes in copulative and cumulative forms (Table 2).

Table 2: Explanatory powers of the non-competitive view and competitive view

	copulative	cumulative	conflictive
Non-competitive view	+	+	−
Competitive view	+	+	+

5. Anticipated counter-arguments and counter-counter-arguments

To our position of the competitive view, the following counter-argument is anticipated: Contrary to the competitive view, conflictive forms are actually often impossible. However strong the attitudes the speaker may hold, the phrases *yookoso* (Welcome), *subarashii wa ne* (It's wonderful, isn't it?), and *sugosoo nante* (oh, you intend to spend) for example cannot help but to be pronounced as their lexical accents direct (e.g. High-Low-Low-Low for *yookoso*, Low-High-High-High-Low for *subarashii*, Low-High-High-Low for *sugosoo*). This indicates that the basic pitch pattern of Standard Japanese is

non-conflictive. This is the content of the counter-argument.

However there are at least two problems in this counter-argument. The first problem is that this counter-argument confuses actuality and possibility. The fact that a phrase is not pronounced in the conflictive form at the present time does not ensure the impossibility of that phrase's being pronounced in the conflictive form. The counter-argument above, taking non-actuality of the conflictive form as a sufficient condition of impossibility of this form, is infected with a logical fallacy. The second problem is that this counter-argument does not take into consideration the variety of speaker "characters," social-psychological types. Henceforth I will explain these two problems in detail.

5.1 Actuality and possibility

Here I will show the counter-argument to the competitive view above contains a logical fallacy: confusing actuality and possibility. Let us take adjective-nai such as *karakunai* (not hot) as an example.

Traditionally adjective-*nai* has been pronounced in the copulative form only (Sound 24, Figure 24) with attitudes of reasking (Did you say it's not hot?) or questioning (Is it hot or not hot?). Recently younger generation Japanese pronounce it in the conflictive form with a strong attitude of requesting agreement (I feel it is hot and I am sure you agree with me. Do you agree with me?) (Sound 25, Figure 25).

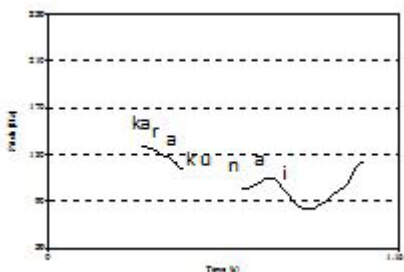**Sound 24: *karakunai?* (It isn't hot?) in copulative form**

Figure 24: Pitch of Sound 24

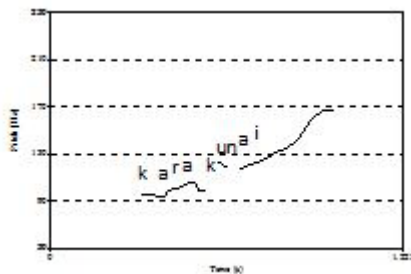**Sound 25: *karakunai?* (It isn't hot?) in conflictive form**

Figure 25: Pitch of Sound 25

The point here is that the conflictive form of adjective-*nai* must have been potentially possible, although the possibility was actualized just recently. Now we all know that it is wrong to argue "However strong an attitude the speaker may hold, phrases of adjective-*nai* cannot help but to be pronounced as their lexical accents direct. This indicates that the basic pitch pattern of Standard Japanese is non-conflictive." The non-existence of a conflictive form at present does not necessarily mean it's impossible.

Then how is adjective-*nai* in the conflictive form activated? This can be understood as an analogy from verbal expressions (Sadanobu 2005a, b). As is

described in Section 4.1, the verbal negative form (verb-*masen*, and verb-*nai* probably) is pronounced in the conflictive form due to strong attitudes, such as strong confirmation or inducement. Adjective-*nai* in the conflictive form connected with strong attitudes cannot be understood unless we admit IS.

5.2 Speaker's characters (social-psychological type)

The second problem of the counter-argument above is that it does not take into consideration the variety of speakers. Indeed it is not rare that a seemingly impossible conflictive form is actualized in accordance with speaker "characters," i.e. social-psychological types (Sadanobu 2010–12, 2011). Before judging a conflictive form as impossible, we must pay attention to the diversity of "*yakuwarigo*" (Kinsui 2003).

Let us take an utterance by Reika Ayanokoji (Photo 1), a famous figure at Universal Studios Japan as an example (Sound 26).
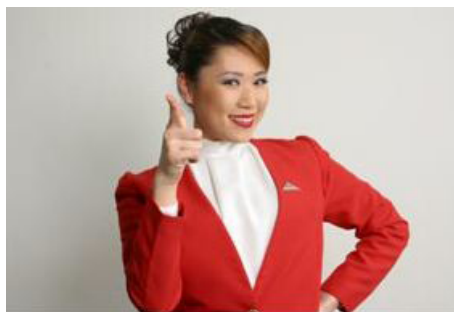

**Photo 1: Reika Ayanokoji at USJ
[http://www.usj.co.jp/CWY/meet_chara/profile02.html, Last retrieved:
2012 Nov. 23.]**

Figure 26: Pitch of Sound 26

Sound 26: Utterance 1
yokoso okoshi
kudasaimashita
(Welcome) by
Reika Ayanokoji
[http://www.yout
ube.com/watch?v
=YXFPGmPUwt4,
Last check: 2012
Nov. 23.

In this utterance, a seemingly impossible conflictive form is actualized. The part of *yokoso* is pronounced with rising pitch and its lexical accent (High-Low-Low-Low) is totally neglected (Figure 26).

Here is another utterance by Reika Ayanokoji, which includes subarashii wa ne (It's wonderful, isn't it?) and sugosoo nante (oh, you intend to spend) that look impossible to pronounce in the conflictive form at first sight, as well as yokoso (Sound 27).

Sound 27: Utterance 2 by Reika Ayanokoji

http://www.youtube.com/watch?v=pgwkatc_Bhw, Last check: 2012 Nov. 23.

And again conflictive forms are actualized with lexical accents neglected in her eccentric and excited rising pitch (Sound 28, Figure 27, and Sound 29, Figure 28). Sound 28: *subarashii wa ne* (It's wonderful, isn't it?) in Sound 27   Figure 27: Pitch of Sound 28

Sound 29: *sugosoo nante* (oh, you intend to spend) in Sound 27

Figure 27: Pitch of Sound 27

Such a way of speaking is not special to the "acted" utterances of Reika Ayanokoji. It is more widely observed in actual utterances by speakers of the hyperactive character.

In summary, the counter-argument against the competitive view lacks validity firstly because it confuses actuality and possibility (5.1), and secondly because it does not take the variety of speaker characters into consideration

(5.2).

6. Conclusions

The conclusions of this paper are as follows.

First, the relationship between lexical accent and phrasal intonation in Standard Japanese can be better understood by the competitive view rather than by the traditional non-competitive view. Lexical accent and phrasal intonation are potentially competing to generate their own pitch form (i.e. conflictive form).

Second, the superiority of the competitive view over non-competitive view is supported by two observations: (i) It is only the competitive view that can explain the conflictive form, which is much more common and widespread than has traditionally been thought; (ii) By adopting Iconicity of Strength (i.e. the idea that the stronger the speaker's attitude is, the more likely the corresponding prosody has strong enough power to generate a conflictive form), the competitive view rightly expects a lack of strong attitudes in copulative and cumulative forms.

Third, Iconicity of Strength can be justified at least in two ways. One is perceptual tests using quasi-minimal pairs, and the other is the possibility of an upset victory of accent over intonation.

And fourth, non-conflictive forms have traditionally been thought of as basic, probably because it is rare for informants to show strong attitudes and prominent characters in the laboratory environment.

In order to deepen our understanding of accent and intonation, we must expand our scope beyond the laboratory wall.

Notes

1: Although a sound wave graph is often presented together with the pitch graph, it is omitted in this paper. This is because this paper includes sound

itself and so bulky graphs of sound waves seem dispensable.

2: The final part *zo* sounds like *do* in some audio and to some listeners. I judged the final part as *zo* mainly for the following two reasons: (i) The manuscript of this data reads *zo* rather than *do*, and (ii) *do* is less common than *zo* (at least for Osaka female speakers).

3: It may seem possible to divide the copulative *nani?* and conflictive *nani?* as two separate homonymic words. This treatment, however, is not adopted in this paper for two reasons. The first reason is that the copulative nani? and conflictive nani? share the interpretation of weak surprise and they are too close to divide, as is indicated by the perceptual test. If we divided them as two separate words it would be difficult to explain this semantic closeness between them. The second reason is that besides the case of *nani?* it is widely attested that a single sentence varies its pitch forms (i.e. conflictive or non-conflictive) in accordance with the strength of attitude. Once we admit this general tendency, the division treatment of *nani?* is not necessary any more.

Acknowledgements

References

Abe, Isamu. 1998 Intonation in Japanese. In Hirst, Daniel, and Albert Di Cristo (eds.), *Intonation Systems: A survey of Twenty Languages*, 360–375, Cambridge, U.K.; New York, N.Y.: Cambridge University Press.

Akinaga, Kazue. 1966 Nihongo no hatsuon: Intoneeshon nado. *Kouza Nihongo Kyouiku*, 2, Institute of Language Teaching, Waseda University, 48–60.

Amanuma, Yasushi, Kazuo Ootsubo, and Osamu Mizutani. 1978 *Nihongo Onseigaku*. Tokyo: Kurosio Publishers.

Fujiwara, Yoichi. 1997 Nihongo Hougen'onno Jissou. Tokyo: Musashinoshoin.

Kawakami, Shin. 1963 Bunmatsu nado no joushouchou ni tsuite. *Kokugo Kenkyu*, 16, Kokugo Kenkyukai at Kokugakuin University, 25–46.

Kinsui, Satoshi. 2003 *Vaacharu Nihongo: Yakuwarigo no Nazo*. Tokyo: Iwanami.

Kori, Shiro. 1989 Kyouchou to intoneeshon. In Miyoko Sugito (ed.) *Nihongo no On'in*, Onsei (Book I), 316–342, Tokyo: Meiji Shoin.

Moriyama, Takuro. 1989 Bun no imi to intoneeshon. In Yutaka Miyaji (ed.) *Nihongogaku Yousetsu*, 172-196, Tokyo: Meiji Shoin.

Sadanobu, Toshiyuki. 2005a Diverse relationships between intonation and accent in Japanese prosody. *Japanese Linguistics*, 17, The National Institute for Japanese Language, 5–25.

Sadanobu, Toshiyuki. 2005b *Sasayaku Koibito, Rikimu Repootaa: Kuchi no Naka no Bunka*. Tokyo: Iwanami.

Sadanobu, Toshiyuki. 2011 *Nihongo Shakai Nozoki Kyarakuri: Kaotsuki, Karadatsuki*, Kotobatsuki. Tokyo: Sanseido.

Sadanobu, Toshiyuki. 2010–2012 An *Unofficial Guide for Japanese Characters*. [http://dictionary.sanseido-publ.co.jp/wp/author/sadanobu-e/ (Sanseido Word-Wise Web)]

Sugito, Miyoko. 2005 On the speech database of three years' prattle and chat by baby Emi and her mother. *Journal of the Phonetic Society of Japan*, 9 (3), 52–57.