

The Prosody of Public Evidence in Japanese: A rating study

Yurie Hara and Shigeto Kawahara
City University of Hong Kong and Rutgers University

1. Introduction

The correlation between Information Structure and prosody has been central in theoretical and experimental studies. In particular, many studies have shown that focus induces particular prosodic patterns cross-linguistically. In Japanese, for instance, pitch register of the words following a focused element is reduced (called ‘post-focus/focal reduction’ in Sugahara 2003 and Ishihara 2003, 2007; see also Tomioka 2009; Kitagawa 2005, 2006). Our study identifies a similar yet distinct kind of deaccentuation which can be regarded as a reflection of another semantic-pragmatic phenomenon, evidentiality.

2. Background

2.1. Biased Question

In Japanese, rising negative questions like (1) express bias meanings which parallel English preposed negative questions (Romero & Han, 2004) or tag questions (Reese, 2007). That is, the question accompanies an implicature that the speaker has a bias toward the positive answer (‘Japanese vegetables are expensive.’). The rising intonation for this construction has two variants. In (1-a) and Figure 1, the lexical accent of *taka’ku* (H*+L) is retained.¹ In (1-b) and Figure 2, the lexical accent is deleted (i.e., deaccented).

- (1) nihon-no yasai, takaku-nai?
Japan-GEN vegetables expensive-NEG
‘Aren’t Japanese vegetables expensive?’
- a. taka’ku nai↑
L%H*+L L%H%
- b. takaku nai↑
%LH- H%

2.2. Deaccentuation

Deaccenting has been widely discussed in the domain of Information Structure (Halliday, 1967). For instance, in many intonation languages (English, German, etc) a post-focus (or post-nuclear) deaccenting takes place after focused word. In Japanese, post-focus/focal reduction (or compression of pitch register) takes place (Sugahara, 2003; Ishihara, 2003, 2007; Tomioka, 2009). In (2), the *wh*-word *nani-o* is in focus and receives the prominent F₀ peak, whereas *nomiya-de* and *nonda* are post-focus-reduced as depicted by the shade in Figure 3.

*The presented research is supported by City University of Hong Kong New Staff Start-up Grant (Project No. 7200192). ‘Experimental investigation into prosodic representation of bias and evidentiality’. We would like to thank Yuki Hirose at University of Tokyo and our research assistants, Yuli Feng and Kenji Ogawara. We are also grateful to Shinichiro Ishihara, Sarah Korostoff, Mana Kobuchi-Phillips, Sunyoung Oh, Melanie Pangilinan, Katsuhiko Sano, Shinichiro Sano, Yasutada Sudo, Jiwon Yun and the audience at WCCFL 29 for their helpful comments.

¹In this paper we use the J_ToBI system for tonal transcription (Venditti, 2005).

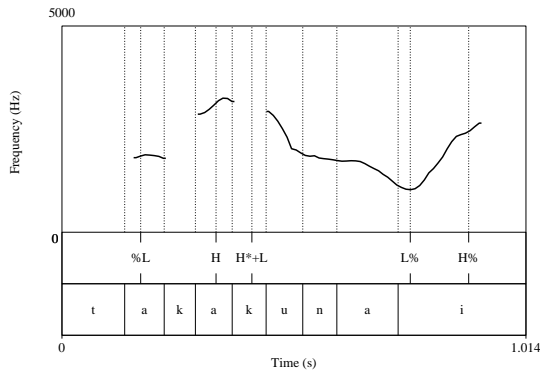


Figure 1: Accented

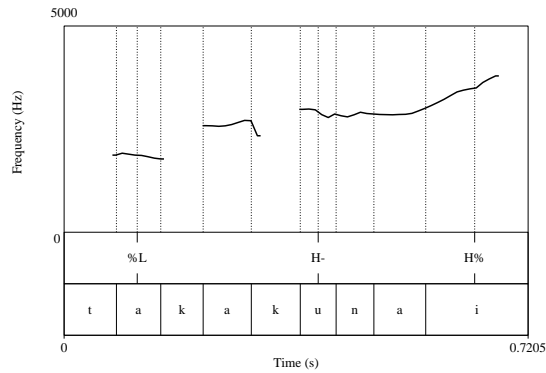


Figure 2: Deaccented

- (2) Naoya-ga nani-o nomiya-de nonda no?
 Naoya-NOM what-ACC bar-LOC drank Q
 ‘What did Naoya drink at the bar?’

(Ishihara, 2003:p. 52)

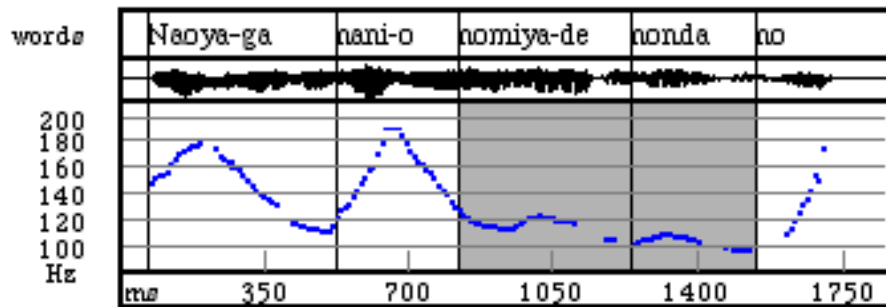


Figure 3: Ishihara, 2003; p. 53

The difference between the deaccenting under the current investigation and the post-focus-reduction is that the lexical accent is completely deleted in the deaccentuation while the F_0 peak is lowered but the lexical contrast of pitch accent is still retained in the post-focus-reduction.² Indeed, we argue that the deaccenting has a discourse function different from the post-focus-reduction. That is, by way of deaccentuation the speaker indicates that he or she has enough evidence for the proposition expressed by the sentence.

2.3. Evidentiality

Evidentiality is “[t]he grammatical encoding of the speaker’s [...] grounds for making a speech act” (Faller, 2002:p. 2). For instance, Japanese has several evidential morphemes. *Soo da* indicates that the speaker has hearsay evidence.

- (3) Ame ga hutteiru soo da.
 rain NOM fall EVID COP
 ‘They say it is raining.’

(Aoki, 1986:p. 230)

²Ishihara (2003) also discusses the difference between deaccentuation and post-focus-reduction. Ishihara (2003) says that no post-lexical operation deletes pitch contrasts; hence the distinction of the lexical accent is still observable after the post-focus-reduction. Ishihara (2003) also notes one exception to his generalization that the *wh*-phrases lose their accents in NPI *mo* constructions (see Kuroda, 2006).

Yoo da and *rasii* mark inferential evidence. The speaker uses *yoo da* when perceptual evidence is available (4).

- (4) Kono kusuri wa yoku kiku yoo da.
this medicine TOP well work EVID COP
'I infer from my own experience that this medicine works well.' (Aoki, 1986:p. 232)

If the evidence available is 'circumstantial or gathered through sources other than one's own senses' (Aoki, 1986:p. 230), *rasii* must be used:

- (5) Kono kusuri wa yoku kiku rasi-i.
this medicine TOP well work EVID-PRES
'I infer from what I heard that this medicine works well.' (Aoki, 1986:p. 232)

The selection of the morpheme depends on the nature of the evidence available in the context. Thus, if the speaker only has indirect evidence, the use of *yoo da* becomes infelicitous:

- (6) #Nonda koto nai kedo, kiita hanashi ni yoruto, kono kusuri wa yoku kiku yoo da.
drank NML NEG but, heard story DAT depend this medicine TOP well work EVID COP
'I have never taken it but according to what I heard, I infer from my own experience that this medicine works well.'

Going back to the Japanese prosodic variation of the adjectives in (1-a) and (1-b), an introspection-based data suggests that the use of deaccentuation seems to correlate with the evidentiality associated with the proposition embedded in the utterance (e.g., Hara & Kawahara, 2008). That is, the speaker uses deaccentuation when the speaker has public evidence stronger than hearsay or circumstantial evidence.³ To illustrate, let us see the following negative question whose predicate is an adjective.

As noted above, the lexical accent of the adjective is either deleted or retained. The deaccentuation is felicitous only when the conversation participants have public evidence for the positive answer as in (7).

- (7) Public Evidence Context
A and B just went to a Japanese supermarket and realized that Japanese vegetables are twice as expensive as local ones. A asks B:
a. #taka'ku nai↑ (Accented)
b. ✓ takaku nai↑ (Deaccented)

In contrast, when there is no public evidence available, the accented adjective is preferred.

- (8) No Public Evidence Context
A has just arrived Hong Kong and B told A that she can get Japanese vegetables from a Japanese supermarket. A asks B.
a. ✓ taka'ku nai↑ (Accented)
b. #takaku nai↑ (Deaccented)

In short, deaccentuation requires a context where the speaker has public evidence which supports the positive answer. Although this deaccentuation pattern has been documented in a sociolinguistic context (Tanaka, 2010)⁴ or pedagogical context (Wakuda, 2003), its semantic property has not been systematically investigated in theoretical linguistics. To objectively confirm the correlation between deaccentuation and availability of public evidence, we conducted an experiment using data elicited from a pool of theoretically unbiased speakers (see Schütze 1996) to test the hypothesis that evidentiality

³In Hara & Kawahara (2008), we suggested that deaccentuation requires direct evidence. As pointed out by Yasutada Sudo (personal communication), however, deaccentuation is available for predicates of personal taste (Lasersohn, 2005) such as *mazui* 'distasteful', and hence direct evidentiality is not appropriate to describe the distributional pattern as one cannot have direct evidence for someone else's personal taste.

⁴For example, our consultants share a strong intuition that the use of deaccentuation is more prominent in women and younger generations, which is in accordance with the findings in Tanaka (2010).

affects the choice of the intonational patterns. The experiment shows that deaccentuation is a marker for public evidentiality.

3. Method

The experiment was a naturalness rating study in which native speakers of Japanese judged the naturalness of different combinations of accentual patterns and degree of public evidence.

3.1. Stimuli

The stimuli had two fully-crossed factors—accentedness and public evidentiality—which resulted in four conditions. Each stimulus consisted of a context, which distinguished evidentiality, as exemplified in (7) and (8).

Each context was followed by a target sentence which contained an adjective predicate with an underlying lexical accent. Each condition had 16 items, resulting in 64 target sentences (16 items * 4 conditions). Each target sentence was pronounced with and without an accent. 64 fillers were included.

3.2. Recording

A native female speaker of Tokyo Japanese pronounced the stimuli at a sound-attenuated room at Research Laboratory for Phonetics and Cognitive Studies of City University of Hong Kong. She produced all the stimuli in isolation, and the stimuli were presented in Japanese orthography printed on paper. For each sentence, the speaker was asked to pronounce it with an accented and a deaccented predicate.

3.3. Procedure

The rating experiment was conducted in the Sound Lab at the University of Tokyo. The stimuli were embedded in an assessment management software, Perception, which was used for the rating study. The participants were asked to take the test in a sound-attenuated room and wear headphones. The first page of the test showed the instructions.

In the main section, the participants were asked to read the context and listen to each stimulus, and then judge the naturalness of stimulus on a 5-point scale rating (provided in Japanese): very natural, somewhat natural, undecidable, somewhat unnatural, very unnatural. They were asked to judge the naturalness of the intonation against the context. They were explicitly asked to pay attention to the naturalness of the speech rather than the content of the stimulus sentence. They were also reminded not to rate the naturalness in terms of the social appropriateness of the speech. This precaution was made since the deaccented rendition can be perceived as a casual speech, and some participants might judge the deaccented sentences as “prescriptively inappropriate”.⁵

The test started with a practice session where the participants ran through five practice items, which were unique to the practice block. The main experiment was organized into four blocks separated by three break signs. Each block contained 16 items. None of the stimuli were repeated and the order of the stimuli within each block was randomized by Perception. No minimal pair sentences appeared next to each other.

3.4. Participants

Fourteen native speakers of Tokyo Japanese participated in the rating experiment. They are undergraduate students recruited from the University of Tokyo and received 1000 Japanese yen for compensation.

⁵The deaccented prosody of adjectives is also considered to be a new usage and is common among young speakers.

3.5. Statistics

The responses were converted to numerical values as follows: very natural=5; somewhat natural=4; undecidable=3; somewhat unnatural=2; very unnatural=1. To analyze the results, a general linear mixed model (Baayen, 2008; Baayen et al., 2008; Bates, 2005) was run using the `lme4` package (Bates et al., 2011) implemented in R (R Development Core Team, 1993–2011). Accentuatedness and evidentiality were the fixed factors. Speakers and items were the random factors. The p-values were calculated by the Markov chain Monte Carlo method using the `LanguageR` package (Baayen, 2009).

If accentuation of the predicate depends on the evidentiality of the context, then the dependency will be expected to result in a significant interaction between accentuatedness and evidentiality.

4. Results

Figure 4 shows the average naturalness ratings in each condition: accented adjectives with public evidence, unaccented adjectives with public evidence, accented adjectives without public evidence, unaccented adjectives without public evidence. In the contexts with public evidence, the speakers judged the predicates without accents more natural. On the other hand, in the contexts without public evidence, the speakers judged the predicates with accents more natural. Because of this asymmetry, the interaction between accentuatedness and evidentiality was significant in the linear mixed model analysis ($t = -24.5, p < .001$). As for the main effects, accentuatedness showed a significant main effect because deaccented predicates are generally preferred (Accented=3.59, Deaccented=3.83, $t = 14.91, p < .001$). Evidentiality also had a significant effect because contexts with public evidence are generally preferred (PublicEvidence=3.73, NoPublicEvidence=3.69, $t = 17.76, p < .001$).

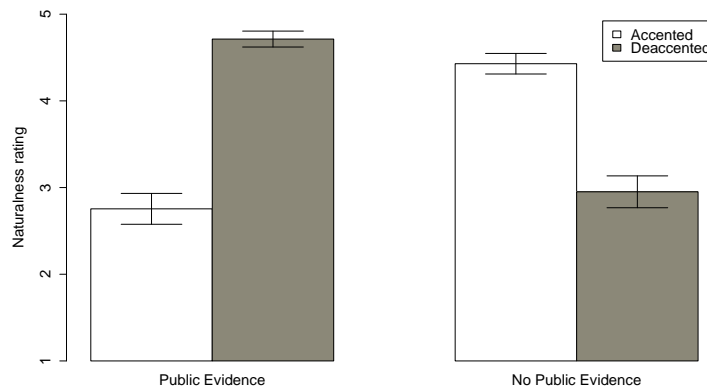


Figure 4: Average Naturalness Ratings

5. Discussion

The result shows that deaccentuation marks the utterance as EVID, which requires a context where the interlocutors have evidence for the embedded proposition (i.e., the positive answer). Hence, deaccentuation is discouraged in no public evidence contexts like (8). In strong evidence contexts like (7), in contrast, deaccentuation is encouraged. Then a question that arises is: Why is a deaccented question ever possible? The proposition is already supported by the public evidence. Why does the speaker still ask? The answer is that a deaccented question is not an information-seeking question but rather a meta-question about the discourse states in the sense of Barker (2009).

5.1. Barker (2009) on English Clarity Assertions

Barker (2009) proposes that an English clarity assertion like (9), which appears to be uninformative, is actually a meta-assertion about the discourse states.

(9) It is clear that Abby is a doctor. (Barker, 2009:p. 253)

According to Barker (2009), the assertion of (9) is paradoxical. In order for (9) to be uttered felicitously, there must be public evidence that Abby is a doctor, e.g., wearing a lab coat and a stethoscope. In contrast, such a requirement is not necessary for the plain assertion like (10). The speaker can truthfully utter (10) with his or her private evidence.

(10) Abby is a doctor. (Barker, 2009:p. 253)

Thus, the clarity statement requires public evidence which makes the truth condition of the sentence obvious to the conversation participants; hence asserting the clarity statement should become uninformative.

In analyzing clarity assertions, Barker (2009) makes use of two levels of context update (Stalnaker, 1998), main effects and side effects (Barker, 2002; Potts, 2005; Shan, 2005). The main effect is the truth-conditional content of the assertion, while the side effect includes conventional implications, presuppositions, the introduction of discourse referents, etc. Now, according to Barker (2009), the main effect of the clarity assertion is trivial, while the side effect is not. The side effect of the clarity assertion is to announce that enough evidence is available to justify concluding the content of the main effect. We can never be 100% sure of any proposition even when we have a good body of evidence. At one point in the conversation, we have to make up our mind and conclude to add the proposition into the common ground. Barker (2009) argues that clarity is a vague and gradable notion and measures degree of justification. The clarity assertion tells the interlocutor that according to the state of the discourse, which includes the information about the standard for clarity, the public evidence meets the standard; hence we should conclude p . Therefore, (9) translates as ‘The publicly available evidence justifies concluding that Abby is a doctor’.

Barker (2009) models the semantics of the clarity assertion using Kratzer’s (1991) semantics of modality. Let f be an epistemic modal base (11-a) and g a stereotypical ordering source (11-b). $\bigcap f(w)$ is the set of worlds that are consistent with the common ground in w and further restricted by the ordering source $g(w)$ as in (11-c). Barker (2009) introduces a function μ that maps worlds to degrees which are ordered according to g as in (11-d). Thus, ‘It is clear that p ’ (\approx ‘The public evidence justifies p ’) is true if and only if in any world w such that w is consistent with the common ground and the degree of the clarity in w exceeds the standard, p is true in w (11-e).

- (11) a. Let W be a set of possible worlds.
 Let $f : W \rightarrow \wp(\wp(W))$ and $f(w)$ represent the common ground in w .
- b. Let $g : W \rightarrow \wp(\wp(W))$ and $g(w)$ represent a set of assumptions about the normal course of events in w .
- c. $u \leq_{g(w)} v$ iff $v \in q$ implies $u \in q$ for any $q \in g(w)$.
- d. Let μ be a degree-valued function on W such that $w \leq_{g(w)} w'$ implies $\mu(w) \leq \mu(w')$ for any $w, w' \in W$.
- e. Let (w, d) be an evaluation point such that w is a possible world and d is a possible discourse.
 Then, $(w, d) \in \llbracket \text{The public evidence justifies } p \rrbracket$ iff $\forall x \in \bigcap f(w). \mu(x) \leq d(\text{clear})$ implies $x \in p$,
 where $d(\text{clear})$ is the operative standard for skepticism at the evaluation point. (adapted from Barker, 2009:p. 268)

This way, Barker’s (2009) analysis escapes the paradox. The side effect of the clarity assertion sets up a cutoff point for justifying the strength of the evidence; hence is informative about the state of the discourse.

5.2. Japanese Deaccented Questions

The deaccented biased questions in Japanese has the same seeming paradox. In a deaccented question p -EVID-*nai*?, p is marked by EVID; hence the public evidence already supports p . The speaker does not need information about the fact of the world. We can straightforwardly extend Barker's analysis on clarity assertions to our evidenced questions. Instead of announcing the justification of the evidence, the speaker expresses an uncertainty about the sufficiency of the evidence available in the discourse, 'Is the evidence clear enough to conclude that p ?'. In order to ask this question, there must be already public evidence available in the discourse. That is, the common ground must already be restricted so that all the worlds in the common ground support that the public evidence is available for the proposition p (12-a). Hence, the deaccentuation (i.e., EVID-marking) is only possible with the public evidence context. Once this presupposition is satisfied, the question meaning inquires whether the clarity of the public evidence exceeds the standard in the discourse. To model a question, a simple Hamblin-style (1973) semantics should suffice here. Thus, the meaning of p -EVID-*nai*? is a set which contains the proposition 'The public evidence justifies that p ' and its complement 'The public evidence does not justify that p ' as in (12-b).

- (12) a. $\llbracket p$ -EVID-*nai*? \rrbracket is defined at (w, d) iff there is public evidence which supports p in any world in $\bigcap f(w)$.
- b. Where defined,
 $\llbracket p$ -EVID-*nai*? \rrbracket
 $= \{ \llbracket \text{The public evidence justifies } p \rrbracket, \llbracket \text{The public evidence does not justifies } p \rrbracket \}$,
where $\llbracket \text{The public evidence does not justify } p \rrbracket = W \times D - \llbracket \text{The public evidence justifies } p \rrbracket$
and D is the set of discourses.

To illustrate, in order to ask the question (1) with deaccentuation, the common ground must already support that there is some evidence that Japanese vegetables are expensive. If so, the side effect of the question partitions the common ground into two blocks which are represented by the proposition 'The public evidence justifies that Japanese vegetables are expensive' and its complement 'The public evidence does not justify that Japanese vegetables are expensive'.

Therefore, the evidenced question does not inquire about the fact of the world but inquires about the epistemic standard for justification of the public evidence for p . This conclusion is independently supported by an experiment conducted by Wakuda (2003), which documents prosodic patterns of the clause containing a negation (i.e., ' p -*nai*') when it is uttered as *hitei-hyoomei* 'asserting negation' and when it is uttered as *dooi-motome* 'seeking for agreement'. Wakuda (2003) concludes that deaccentuation takes place more often when the expression functions as 'seeking for agreement' than when it functions as 'asserting negation'. See Tanaka (2010) who reaches the same conclusion.

In short, the biased question which contains deaccentuation presupposes that public evidence for the positive answer is available. The main effect of the question is superfluous as the evidence which supports that the positive answer is already available, while the side effect of the question is informative because the speaker is uncertain whether the clarity of the available evidence is strong enough compared to the discourse standard.

6. Concluding Remarks

6.1. Summary

Deaccentuation marks an embedded proposition in a biased question as evidenced, and the utterance results in a meta-question about the discourse states. The speaker asks whether the available evidence is clear enough to support the positive answer. Our experiment thus shows that intonation can mark evidentiality as well as focus structures.

6.2. Remaining Issues

There are several future directions pertaining to our result. First, is it possible to connect the deaccentuation as public evidentiality and the deaccentuation as givenness as discussed in the focus literature? Can we argue that one is derived from the other?

The second issue concerns with syntactic restrictions on deaccenting. We have so far been focusing on adjectives, but there seem to be some interesting restrictions on the kinds of predicates that can undergo deaccentuation. Generally, verbs cannot be deaccented as shown in (13).

- (13) ano hito hashira nai?
 that person run NEG
 'Doesn't that person run?'
 a. ✓hashira' nai↑ (Accented)
 b. *hashira nai↑ (Deaccented)

This restriction may arise from the semantic property of adjective, i.e., vagueness and gradability. Interpreting adjectives always requires the standard in the discourse context. However, Japanese also has a class of modifiers similar to adjectives, sometimes known as *adjectival nouns* (e.g., *kirei* 'pretty' and *sinsetsu* 'kind'). The adjectival nouns are similar to adjectives since they modify nouns and they can be modified by an intensifier adverb like *totemo* 'very', and hence they are gradable predicates. These predicates cannot be deaccented, as in (14).⁶

- (14) ano hito kirei ja nai?
 that person pretty COP NEG
 'Isn't that person pretty?'
 a. ✓kirei' ja nai↑ (Accented)
 b. *kirei ja nai↑ (Deaccented)

One possible explanation is that adjectival nouns are morphologically more complex than adjectives. They have to be followed by an inflecting copula *da*, whose negative form, *ja*, is a contracted form of *de-wa* (copula plus topic).⁷ Hence, we might be able to maintain the gradability condition of the deaccentuation while ruling out the deaccentuation of the adjectival nouns due to their morphological complexity. Follow-up experiments will reveal whether any phonological, morphological or syntactic features of these items interfere with the availability of the deaccentuation.

Last but not least, it is always a fruitful investigation to make the semantic and pragmatic computation of the deaccentuation compositional. Is EVID fully integrated in the syntactic structure or shall we treat it as a floating morpheme? Thus, the Japanese deaccentuation pattern should be further investigated from different perspectives.

References

- Aoki, Haruo (1986). Evidentials in Japanese. Chafe, Wallace & Johanna Nichols (eds.), *Evidentiality: The Linguistic Coding of Epistemology*, Ablex Publishing Corporation, Norwood, New Jersey.
- Baayen, Harald R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press, Cambridge.
- Baayen, Harald R. (2009). LanguageR. R package.
- Baayen, Harald R., Doug.J. Davidson & Douglas. M. Bates (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.
- Barker, Chris (2002). The dynamics of vagueness. *Linguistics and Philosophy* 25, 1–36.
- Barker, Chris (2009). Clarity and the grammar of skepticism. *Mind & Language* 24, 253–273.
- Bates, Douglas (2005). Fitting linear mixed models in R. *R News* 5, 27–30.
- Bates, Douglas, Martin Maechler & Ben Bolker (2011). lme4: Linear mixed-effects models using S4 classes. R package.
- Faller, Martina (2002). *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. Ph.D. thesis, Stanford University.
- Halliday, M.A.K. (1967). Notes on transitivity and theme in English, Part II. *Journal of Linguistics* 3, 199–244.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language* 10, 41–53.
- Hara, Yurie & Shigeto Kawahara (2008). Daccenting, MAXIMIZE PRESUPPOSITION and evidential scale. Barbosa, P. A., S. Madureira & C. Reis (eds.), *Proceedings of the Speech Prosody 2008 Conference*, Editora RG/CNPq, Campinas, Brazil, 509–512

⁶Deaccentuation of verbs and adjectival nouns is sporadically observed, as documented in Tanaka (2010).

⁷See Shibatani (1990) and Tsujimura (2007) for the discussion of the adjectival and nominal features of the adjectival nouns.

- Ishihara, Shinichiro (2003). *Intonation and interface condition*. Ph.D. thesis, Massachusetts Institute of Technology.
- Ishihara, Shinichiro (2007). Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24, 137–167.
- Kitagawa, Yoshihisa (2005). Prosody, syntax and pragmatics of wh-questions in Japanese. *English Linguistics* 22:2, 302–346.
- Kitagawa, Yoshihisa (2006). Wh-scope puzzles. *NELS35: Proceedings from the 35th North Eastern Linguistics Society*, GLSA, University of Massachusetts, Amherst, Amherst, Massachusetts.
- Kratzer, Angelika (1991). Modality. von Stechow, A. & D. Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, de Gruyter, Berlin, 639–650.
- Kuroda, S.-Y. (2006). Prosody and the syntax of indeterminates. *Indiana University Working Papers in Linguistics*, IULC Publications, vol. 5.
- Lasersohn, Peter (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28, 643–686.
- Potts, Christopher (2005). *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics, Oxford University Press, Oxford. [Revised 2003 UC Santa Cruz PhD thesis].
- R Development Core Team (1993–2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reese, Brian (2007). *Bias in Questions*. Ph.D. thesis, University of Texas at Austin.
- Romero, Maribel & Chung-Hye. Han (2004). On negative yes/no questions. *Linguistics and Philosophy* 27:5, 609–658.
- Schütze, Carson (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. University of Chicago Press, Chicago.
- Shan, Chung-chieh (2005). *Linguistic Side Effects*. Ph.D. thesis, Harvard.
- Shibatani, Masayoshi (1990). *The Languages of Japan*. Cambridge.
- Stalnaker, Robert (1998). On the representation of context. *Journal of Logic, Language, and Information* 7, 3–19.
- Sugahara, Mariko (2003). *Downtrends and Post-FOCUS Intonation in Tokyo Japanese*. Ph.D. thesis, University of Massachusetts, Amherst.
- Tanaka, Yurkari (2010). *Shutoken ni Okeru Gengo Doutai no Kenkyuu*. Kasama Shoin.
- Tomioka, Satoshi (2009). Why questions, presuppositions, and intervention effects. *Journal of East Asian Linguistics* 18, 253–271, URL <http://dx.doi.org/10.1007/s10831-009-9053-0>. 10.1007/s10831-009-9053-0.
- Tsujimura, Natsuko (2007). *Introduction to Japanese Linguistics*. Blackwell, 2nd edn.
- Venditti, Jennifer J. (2005). The J-ToBI model of Japanese intonation. Jun, Sun-Ah (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford University Press, 172–200.
- Wakuda, Miho (2003). I-keiyoshi+nai no inritsuteki tokucho: akusento intoneshon setsuzokujikan no sokumen kara. *Waseda Daigaku Nihongo Kyouiku Kenkyuu*, Waseda University, 125–139.