On the (lack of) correspondence between syntactic clauses and intonational phrases

Shinichiro Ishihara

Preprint (version 2.6), March, 2021

Abstract

This paper discusses the syntax–prosody mapping at the clause-level, or the potential absence thereof. In particular, the paper reexamines the validity and the necessity of the clause-level mapping constraint, MatchClause, proposed in Match Theory (Selkirk 2009, 2011). First, it will be shown, based on the results of previous empirical studies, that data from embedded clauses in Japanese do not provide positive evidence for MatchClause. The discussion includes a critical review of Selkirk's (2009) analysis of *wh*-prosody in Fukuoka Japanese. Secondly, it will be claimed, following other recent proposals in the literature, clausehood is irrelevant in what has been assumed to be mapping between so-called illocutionary clauses to intonational phrases (ι). An alternative approach is suggested in which the intonational phrase (ι) is related to speech acts rather than the syntactic clausehood.

Keywords

syntax–prosody interface, Match Theory, Japanese, embedded clauses, speech act

1 Introduction

This paper discusses the syntax–prosody mapping at the clause-level, or the potential absence thereof. Recently proposed theories of intonation (in particular, Itô and Mester 2007, 2012, 2013 and Selkirk 2009, 2011) assume a correspondence principle between syntactic *clauses* and prosodic constituents commonly referred to as the *intonational phrase* (ι hereafter). While the syntax–prosody mapping at the clause-level appears theoretically well-motivated and attractive in its simplicity, it is still empirically much less frequently examined in the literature compared to the word and phrase-level mapping. Furthermore, it has been reported for several languages that embedded clauses do not show consistent correspondence with ι's. Although the lack of correspondence between embedded clauses and ι's

can be accounted for, and is even expected, within the current Match Theory, it still seems a legitimate question whether MatchClause is needed in the theory of the syntax–prosody mapping. With this question in mind, this paper reexamines the validity and the necessity of the clause-level mapping principle from both empirical and theoretical perspectives, based on some of the data currently available in the literature. In particular, the paper focuses on two aspects of MatchClause, namely, mapping of embedded clauses onto ι's, and the relevance of illocutionary force (or speech act associated with it) to ι-mapping.

The paper is organized as follows. §2 briefly presents some background information. §2.1 introduces the Syntax–Prosody Mapping Hypothesis (SPMH) put forward by Itô and Mester (2007, 2012, 2013) and Selkirk (2009, 2011) and, in particular, Selkirk's (2011) analysis of the clause-level syntax–prosody mapping. §2.2 presents the assumption made in this paper regarding the status of ι in the prosodic hierarchy of Japanese.

In §3, previous studies on the prosody of Japanese embedded clauses will be reviewed. It will be shown that the majority of these studies cannot be considered evidence for the clause-level mapping (§3.1). Also, Selkirk's (2009) analysis of *wh*-prosody in Fukuoka Japanese will be discussed in detail (§3.2). It will be argued that her proposal cannot be maintained due to empirical problems. Since additional data go against the prediction of the analysis, the arguments for the clause-level mapping made by Selkirk (2009) are no longer tenable. Based on the discussion in §3.1 and §3.2, it will be concluded that Japanese embedded clause data do not provide positive evidence for MatchClause. In addition, §3.3 suggests, albeit tentatively, a possibility to derive apparent mapping of embedded clauses to ι's found in other languages without MatchClause.

§4 discusses the relevance of speech act to ι-mapping as well as the irrelevance of clausehood. §4.1 discusses cases in which clauses do not map to ι, as well as cases where non-clauses map to ι just like clauses. In particular, it will be proposed (§4.2), along with previously made proposals on root clauses (Downing 1970; Selkirk 2005, 2009, 2011), parentheticals (Potts 2005; Selkirk 2005; Dehé 2014; Güneş 2014, 2015; Truckenbrodt 2015) and certain types of topics (Ebert 2009; Bianchi and Frascarelli 2010), that a constituent that performs a speech act, be it a clause or a phrase, is mapped to an ι (§4.3). Building on this proposal, an alternative architecture of grammar will be suggested, in which ι is no longer relevant for the syntax–prosody mapping but is instead responsible for the discourse–prosody mapping (§4.4). This model implies a new perspective on the interfaces between prosody and other modules of grammar: only the phonological word (ω) and the phonological phrase (φ) are responsible for syntax–prosody mapping, as proposed in previous theories, while ι is responsible for the discourse/pragmatics–prosody mapping.

§5 summarizes the discussion, as well as raises one fundamental theoretical question for future re-

search regarding the clause-level syntax–prosody mapping.

2 Background

2.1 The Syntax–Prosody Mapping Hypothesis

There are two recent theories of the prosodic hierarchy and the syntax–prosody mapping which put forward a strong version of the *Syntax–Prosody Mapping Hypothesis (SPMH)*. The first one is the *Prosodic Adjunction Theory* by Itô and Mester (2007; 2012; 2013), and the other is *Match Theory* proposed by Selkirk (2009; 2011). They both assume three distinctive categories in the prosodic hierarchy, namely, the phonological word (ω), the phonological phrase (φ), and the intonational phrase (ι, a.k.a. *phonological clause*[1]). The strong version of the SPMH advocated in these theories states that these three prosodic categories systematically, and language-universally, correspond to three syntactic categories, words, phrases, and clauses. (1) lists the syntax–prosody mapping principles assumed in Match Theory.

(1) Syntactic-prosodic constituency correspondence in Match Theory of (Selkirk 2011:439)

a. Match clause

A clause in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ι, in phonological representation.

b. Match phrase

A phrase in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it φ, in phonological representation.

c. Match word

A word in syntactic constituent structure must be matched by a corresponding prosodic constituent, call it ω, in phonological representation.

Among these three levels of syntax–prosody correspondence, the phrase-level correspondence in (1b) has been widely attested cross-linguistically and discussed in the literature since the earliest analyses of the syntax–prosody interface, such as relation-based mapping (Nespor and Vogel 1986), end-based mapping (Selkirk 1984, 1986; Chen 1987), and arboreal mapping (Zec and Inkelas 1990; Inkelas and Zec 1995). Setting aside the independent question of whether another category, *Clitic Group*,

[1]Instead of the intonational phrase, the term *phonological clause* is proposed by Itô and Mester (2013). The reason behind it is two-fold: to make the syntax–prosody correspondence put forward in the SPMH more explicit, and to keep the names of prosodic categories phonetically neutral, as prosodic constituents at this level may be marked by non-tonal phenomena, such as initial strengthening or final lengthening (Junko Ito, p.c.). In the remainder of the paper, I remain neutral as to which term to be used, and use the symbol ι to refer to this level in the prosodic hierarchy (together with ω and φ for the two lower levels).

should be maintained in the universal prosodic hierarchy (Vogel 2009, 2020), there is also abundant evidence for the mapping relation between syntactic and phonological words as well (see, e.g., studies in Grijzenhout and Kabak 2009 for recent relevant discussions).

According to Selkirk (2011), the clause-level correspondence is sensitive to clause types. The motivation for the distinction between clause types comes from an observation in the literature (Downing 1970; Nespor and Vogel 1986; Ladd 1986; Selkirk 2005; Dehé 2009) that certain types of clauses show more consistent mapping to an ι than other types of clauses (see also Güneş 2014, 2015; Truckenbrodt 2015 for more recent discussion). Selkirk calls the former *illocutionary clauses* and the latter *standard clauses*. The standard clause is defined as "the constituent that is the complement of the functional head Comp^0," and "may be syntactically embedded, whether as a complement to a verbal or nominal head, or as a restrictive relative clause within determiner phrase, or in other positions" (Selkirk 2011:452). The illocutionary clause, on the other hand, is "the highest syntactic projection of the sentence and carries its illocutionary force" (Selkirk 2011:452), which Selkirk assumes to be the complement of Force Phrase (ForceP, Rizzi 1997). With the common assumption that only the root (topmost) sentence can carry illocutionary force, ForceP (illocutionary clause) cannot be embedded.[2]

Selkirk (2009, 2011) actually mentions the possibility of embedded ForceP as well, to account for cases of non-root clauses that obligatorily map to ι. She assumes, adopting Potts's (2005) analysis of parenthetical expressions, that "a Force Phrase is syntactically embedded in the case of parentheticals, nonrestrictive relatives and so on, and that it triggers the multi-tier semantic interpretation Potts proposes which gives rise to multiple speech acts" (Selkirk 2009:49). In §4 we will discuss more in detail the specific syntactic constructions mentioned here in relation to illocutionary clauses and their relations to multiple speech acts. For now, the crucial assumptions are that there are two types of clauses, illocutionary and standard clauses, and that they differ in terms of the syntax–prosody mapping: while illocutionary clauses consistently map to ι's, standard clauses may or may not map to ι's.

In order to capture the difference between the two clause types, Selkirk (2009, 2011) proposes that there are two different versions of MatchClause constraints, namely, the general version Match(clause, ι) and the more specific version Match(illocutionary clause, ι). These two versions of MatchClause may be ranked differently with respect to some constraint that masks the effect of a lower-ranked Match constraint. If there is a ranking in which Match(illocutionary clause, ι) dominates a match-blocking prosodic wellformedness constraint (PWC) which, in turn, dominates Match(clause, ι), it is expected that only illocutionary clauses exhibit mapping to ι while standard clauses do not map to ι. In con-

[2]Hamlaoui and Szendrői (2015, 2017) propose an alternative approach to ι-mapping. Instead of specifying particular syntactic nodes (such as TP or CP) as the target of the clause-level syntax–prosody mapping, they propose that it is the "highest verbal projection whose head is overtly filled by the root verb or verbal material" (Hamlaoui and Szendrői 2015:82) that maps to ι.

trast, if MATCH(clause, ι) is ranked above the PWC, together with MATCH(illocutionary clause, ι), then ι-mapping is expected with all clauses. Selkirk (2009) proposes that Fukuoka Japanese is an example of the latter case.[3] It is also suggested (Selkirk 2011:453) that MATCH(illocutionary clause, ι) is universally ranked above MATCH(clause, ι), which results in the universal tendency to show more consistent ι-mapping with illocutionary clauses than with standard clauses.

Under this analysis, the potential lack of mapping between (standard) clauses and ι is explained by the constraint ranking in which MATCH(clause, ι) is outranked by another constraint that masks the matching effect. Therefore, the absence of clause-level mapping in a language does not necessarily constitute a counterargument against MATCHCLAUSE. At the same time, however, the lack of positive evidence also seems to suggest that an alternative approach would be equally worth taking into consideration, which assumes no clause-level syntax–prosody mapping.

2.2 The ι in Japanese

Before going into the main discussion, it is necessary to make an explicit assumption that the level of ι exists in Japanese, as the question of whether this level exists in the prosodic hierarchy of Japanese or not is not entirely a settled issue.

Earlier analyses of Japanese intonation all assumed two distinct levels that correspond to (subparts of) the level of phonological phrase in other languages. These two levels have been referred to with different terms: *minor phrase* and *major phrase* (McCawley 1965; Poser 1984; Kubozono 1993), *accentual phrase* and *intermediate phrase* (Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988), *accentual phrase* and *intonation phrase* (Venditti 2005; Maekawa et al. 2002; Venditti et al. 2008; Igarashi 2015), *accent component* and *phrase component* (Fujisaki and Sudo 1971; Fujisaki and Hirose 1984). Itô and Mester (2012), however, convincingly demonstrate that these two levels can be considered to belong to a single category, φ, with the minor phrase defined as the minimal projection of φ, and the major phrase as all instances of φ.[4]

As for the level of ι, Pierrehumbert and Beckman (1988) claimed that there is no empirical evidence that supports the existence of the level of ι, and hence adopted a hierarchy lacking this level. Japanese ToBI models, both the original J_ToBI (Venditti 2005) and the extended version, i.e., X-JToBI, (Maekawa et al. 2002), adopted the prosodic hierarchy proposed by Pierrehumbert and Beckman (1988).[5]

Kawahara and Shinya (2008), however, reported, based on the results of their production experi-

[3] In §3.2, it will be claimed that this proposal is not tenable on empirical grounds.

[4] See Ishihara (2015:570–576) for a summary of terminologies and the phonological properties of φ in Japanese.

[5] However, for the level that is referred to as "intermediate phrase" by Pierrehumbert and Beckman (1988), the ToBI models adopt the term *intonation phrase*. This means that the intonation phrase within the ToBI models corresponds to the level of φ, not ι, according to the prosodic hierarchy adopted in this paper.

ment, that there are phonetic cues that neither belong to φ nor the utterance (υ), and claimed accordingly that Japanese does have the level of ι. Kawahara and Shinya (2008) examined the intonation of coordinated (main) clauses and corresponding gapping sentences, as in (2), and found that each of the coordinated clauses is realized as an ι in the prosodic structure. Gapping sentences are derived by removing the verbs in brackets.

(2) [$_{\text{Clause}}$ Murasugi-wa namauni-o (moritsuke)], [$_{\text{Clause}}$ Munakata-wa mamemochi-o (moritsuke)], [$_{\text{Clause}}$ Morimura-wa aemono-o moritsuketa].
Murasugi-TOP sea.urchin-ACC put.on.dish Munankata-TOP bean.rice.cake-ACC put.on.dish Morimura-TOP mixed.salad-ACC put.on.dish
'Murasugi put a sea urchin on a dish, Munakata (put) bean rice cake on a dish, and Morimura (put) a mixed salad on a dish.'

(Kawahara and Shinya 2008:64)

According to Kawahara and Shinya (2008), the ι in Japanese can be distinguished from φ by 1) final lowering, 2) creakiness and a pause in final position, and 3) a distinctively large initial F_0-rise as well as pitch reset in initial position. It is also distinguished from υ by the lack of a phrasal H-tone at the right edge.

Note that their finding is in accordance with the prediction of Match Theory. MatchClause constraint calls for the mapping of coordinated clauses onto ι's in the prosodic structures. In the following discussion, I will assume that ι is a distinct level in the prosodic hierarchy of Japanese. In addition, I will assume, following Itô and Mester (2012), that utterance (υ) is not a distinctive category, but a subcategory within the level of ι, namely, the maximal projection of ι. In sum, the prosodic hierarchy assumed in this paper is illustrated in (3).

(3) Prosodic hierarchy of Japanese

- ι the intonational phrase/phonological clause ($\iota^{\max}$ = utterance, υ)
- φ the phonological phrase ($\varphi^{\min}$ = minor phrase)
- ω the phonological word

3 (Lack of) clause-level mapping in Japanese

The main purpose of this section is to examine whether there is enough empirical motivation for MatchClause based on Japanese data. As the data in this section all involve either relative clauses or clausal objects (without illocutionary force), the discussion here concerns the general version of MatchClause. It will be concluded that Japanese does not provide empirical support for Match(clause, ι). The other,

more specific version of MatchClause, Match(illocutionary clause, ι), will be discussed in §4.

§3.1 reviews some of the relevant studies on the prosody of embedded clauses in Japanese. These studies do not distinguish between phrase-level and clause-level mapping. This is presumably because it has long been widely accepted that Japanese does not have the level of ι, as mentioned in §2.2. There are also a few studies that unambiguously indicate that embedded clauses do not behave differently from phrases. Even though Kawahara and Shinya (2008) have found empirical evidence for the presence of ι in Japanese, none of the other studies reviewed here can be considered evidence for the clause-level mapping.

In §3.2, we will discuss in detail the proposal by Selkirk (2009). Based on the prosody of *wh*-questions in Fukuoka Japanese, Selkirk argues for the clause-level mapping. It will be shown below, however, that there are several problems in the empirical claims made by Selkirk (2009). Once these problems are taken into consideration, the argument will no longer hold.

In §3.3, a few comments will be made on the studies on languages other than Japanese (in particular, German, Truckenbrodt 2005; Truckenbrodt and Darcy 2010; Schubö 2020) which present empirical evidence for the mapping of embedded clauses to ι's. It will be pointed out that even those cases where embedded clauses are reported to map to ι's, it is not entirely clear whether it is the clausehood of syntactic constituents that triggers the mapping, or other potential factors, such as length, that are responsible for the mapping.

3.1 Previous studies on the prosody of embedded clauses in Japanese

3.1.1 Relative clauses

There are studies that show that the beginning of relative clauses in Japanese are marked prosodically by a large F_0-rise on the clause-initial word, which is, according to Kawahara and Shinya (2008), one of the phonetic cues for ι. One of the earliest studies on relative clauses was conducted by Uyeno et al. (1979). They compared the pitch contours of sentences like (4a) and (4b), where the left edge of the relative clause (indicated by $[_{\text{RC}}$) is varied.

(4) a. [[$_{\text{RC}}$ ototoi koronda] otona-ga waratta]
day.before.yesterday fell adult-NOM laughed
'The adult [who fell the day before yesterday] laughed.'

b. [ototoi [$_{\text{RC}}$ koronda] otona-ga waratta]
'The adult [who fell] laughed the day before yesterday.'

(Uyeno et al. 1979:184)

They found a high F_0-rise at the beginning of the utterance in (4a), followed by a gradually declining pitch contour, and in (4b), an F_0-rise at the beginning of the relative clause in addition to the utterance-initial one. Uyeno et al. (1981) also found that pauses indicate clause-initial boundaries as well.

Kubozono (1993) compared the four different phrase structures in (5) in order to investigate the correlation between phrase structure and prosodic structure. The results showed that the second word of (5d) showed a larger F_0-rise compared to the second words in the other examples.

(5) a. [[Naoko-no ani-no] [aoi erimaki]]
N.-GEN brother-GEN blue muffler
'Naoko's brother's blue muffler'

b. [Mariko-no [ookina [aoi erimaki]]]
M.-GEN big blue muffler
'Mariko's big blue muffler'

c. [[Ayako-no [men-no erimaki-no]] iromoyoo]
A.-GEN cotton-GEN muffler-GEN design
'design of Ayako's cotton muffler'

d. [aoi [[Yumiko-ga anda] erimaki]]
blue Y.-NOM knitted muffler
'the blue muffler Yumiko knitted'

(adapted from Kubozono 1993:211)

The purpose of his experiment was, however, to examine how the different syntactic branching structures are realized prosodically, rather than to compare relative clauses and other phrases. (5d) is used as an example of the phrase structure that has two phrase-boundaries (at the onset of the second word). Also, Kubozono (1993) does not assume the level of ι. His conclusion from this experiment is that the more syntactic (left) edges appear at a certain position, the higher F_0-rise (which he calls *metrical boost*) results.

While it seems clear from these studies that relative clauses tend to show a large F_0-rise at its onset, it is difficult, if not impossible, to draw a definitive conclusion to the relevant question here, namely, whether these findings can be considered evidence for the clause-level syntax–prosody mapping, because no distinction was made between ι and φ in their analyses of the results.

In addition, there is a study that shows opposite results. Hirayama and Hwang (2019) show that relative clauses undergo downstep triggered by the lexical accent of the preceding phrase, indicating that the relative clause and the preceding phrase are contained in a single φ. Downstep is a phonologically conditioned F_0-downtrend triggered by lexical pitch accents. The domain of downstep has

been assumed to be φ (Poser 1984; Pierrehumbert and Beckman 1988; Kubozono 1993; Ishihara 2016). At the onset of a φ, downstep is cancelled and a pitch reset results. One of the pairs of phrases used in Hirayama and Hwang's (2019) study is shown in (6). The location of F_0-falls due to lexical H*+L pitch accents is indicated with the symbol ¬.

(6) a. [RC1 mayo¬-u] [RC2 **naya¬m-u**] mago
get.lost-NPST worry-NPST grandchild
'(my) grandchild who [is lost] and that worries'

b. [RC1 manab-u] [RC2 **naya¬m-u**] mago
study-NPST worry-NPST grandchold
'(my) grandchild who studies and that worries'

(Hirayama and Hwang 2019:2852, emphasis and brackets added by S.I.)

In each of the phrases, there are two relative clauses that contain one word each (indicated as RC1 and RC2, respectively). (6a) and (6b) only differ in terms of the lexical accent of the first word (which constitutes RC1): the first word of (6a), *mayo¬-u*, is lexically accented, while that of (6b), *manab-u*, is unaccented. The results show that the F_0-peak of the second word (which constitutes the target relative clause, RC2) is consistently lower in (6a) than in (6b). This difference in F_0 is a result of downstep. Given that downstep is assumed to take place within φ, the first two words are contained in a single φ. It can therefore be concluded that there is no ι-boundary between RC1 and RC2.

Even though their results unambiguously show that the relative clauses in their stimuli did not map to an ι, however, there is a potential confounding factor that might be responsible for the lack of clause-level mapping. Since the relative clauses tested in Hirayama and Hwang's (2019) experiment (both RC1 and RC2) contain only one word, it is plausible that RC1 and RC2 were phrased together due to a prosodic wellformedness constraint on the size of prosodic constituents (often referred to as minimality or binarity effects), which would avoid a φ containing only one ω, and instead group two ω's into one φ, violating MatchPhrase.[6] Further research is needed to disentangle the effects of the clause-level mapping from the potential confounding effect of minimality/binarity. It is nevertheless clear that their results cannot be considered supporting evidence for the clause-level mapping.

3.1.2 Clausal objects

When it comes to clausal objects, there seems to be no empirical confirmation that they consistently correspond to an ι in Japanese. Ishihara (2020) compared sentences with different numbers of phrase-boundaries in front of a target word and sentences with a clause boundary at the same position, as

[6] There are studies showing that Japanese is subject to binarity effects, especially in terms of maximality: a φ containing more than three ω's tends to be divided into smaller φ's containing two or three ω's each (Selkirk and Tateishi 1988; Kubozono 1993; Shinya et al. 2004).

shown in (7) (where the second word is the target word, marked in **boldface**).[7] This study was, unlike the previous studies mentioned above, specifically designed to examine possible phonetic differences between phrase and clause boundaries.

(7) a. No XP boundary (0xp)

[$_{DP}$ Yuuta-to **Naoya-wa**] [$_{VP}$ imooto-o paatii-ni maneita]
Y.-and N.-TOP sister-ACC party-to invited
'Yuta and Naoya invited their sisters to the party.'

b. One XP boundary (1xp)

Yuuta-wa [$_{VP}$ **Naoya-o** [$_{DP}$ imooto-no paatii-ni] maneita]
Y.-TOP N.-ACC sister-GEN party-to invited
'Yuta invited Naoya to his sister's party.'

c. Two XP boundaries (2xp)

Yuuta-wa [$_{VP}$ [$_{DP}$ **Naoya-no** imooto-o] paatii-ni maneita]
Y.-TOP N.-GEN sister-ACC party-to invited
'Yuta invited Naoya's sister to the party.'

d. Clause boundary (cp)

Yuuta-wa [$_{CP}$ **Naoya-ga** imooto-o paatii-ni maneita to] omotteita
Y.-TOP N.-NOM sister-ACC party-to invited that was.thinking
'Yuta thought that Naoya invited his sister to the party.'

The results from 14 speakers, as summarized in Figure 1, show that the F_0-peak of the target word is higher when it is preceded by an XP-boundary (1xp), and that it is even higher when preceded by two boundaries (2xp), replicating Kubozono's (1993) results. The results further show that clause boundaries are not necessarily marked with a distinctively higher F_0 (cp) than phrase boundaries (1xp, 2xp). There was no indication that pauses distinguish between phrase- and clause-boundaries, either.

There is another study that shows that the left edge of an embedded clause does not necessarily coincide with a strong F_0-rise that would be expected at the edge of φ or ι. Hirayama and Hwang (2016) examined the prosody of embedded clauses containing a single content word followed by a complementizer, as in (8a), and found that the lexical pitch accent in the matrix subject triggers downstep on the sole content word in the embedded clause, indicating that these two noun phrases are contained in a single φ.[8] The results were comparable to those obtained from stimuli like (8b), in which the object is a noun phrase rather than a clause.

[7] For current purposes, single word (i.e., non-branching) DPs like *[Naoya-o]* in (7b) are not included in the count of XP boundaries.

[8] Ishihara (2020) also reports that the embedded clause subject, e.g., *Naoya-ga* in (7d) shows downstep relative to the preceding matrix clause subject in some of the speakers' data.

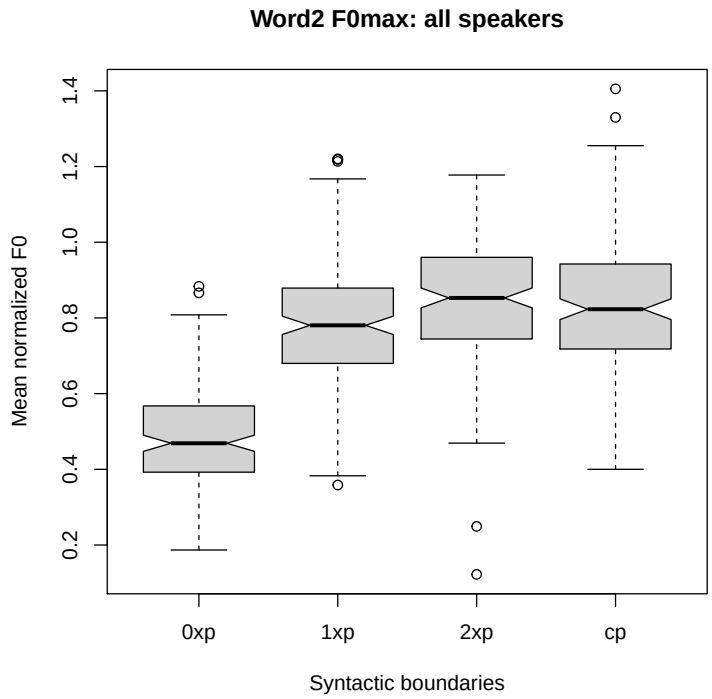


Figure 1: Normalized F_0-maximum on the target word in Ishihara's (2020) stimuli in (7d), based on the data from 14 speakers

(8) a. [ani-wa [$_{CP}$ hana-to] itta.]
brother-TOP flower-COMP said
'My brother said flower.'

b. [ani-wa hana-o utta.]
brother-TOP flower-ACC sold
'My brother sold flowers.' (adapted from Hirayama and Hwang 2016:96)

The results in this study is, as those of Hirayama and Hwang's (2019) study on relative clauses, might be affected by minimality/binarity. Nevertheless, both Ishihara's (2020) and Hirayama and Hwang's (2016) studies do not provide any evidence that embedded clause boundaries are marked by an ι-boundary.

As mentioned in §2.1, the lack of clause–ι correspondence reported in these studies does not serve as counterargument for MatchClause, as Match Theory has an account for the absence of correspondence between embedded (standard) clauses and ι. However, as far as (Tokyo) Japanese is concerned, none of the previous studies provides strong support for MatchClause, either. The only case that seems to provide positive support for MatchClause would be Kawahara and Shinya (2008). Their data, however, do not make any further predictions as to whether the clause-level mapping applies to syntactically embedded clauses, such as relative clauses and clausal objects of main clause verbs, as their data do not involve embedded clauses. Furthermore, there is another potential factor that might

be influencing Kawahara and Shinya's (2008) data, namely, coordination. Coordinated structures have often been used to study the prosodic realization of syntactic branching structures (Ladd 1986, 1988; Wagner 2005; Kentner and Féry 2013; Truckenbrodt and Féry 2015). Many of these studies show that differences in the coordinated structure are realized prosodically by differentiating boundary strengths. If a similar disambiguation effect is present in Japanese, it might be the case that the ι-boundaries found at clause boundaries in Kawahara and Shinya's (2008) data are motivated by the coordinated structure, rather than the syntactic status as a clause, in order to disambiguate the clause-internal syntactic structures and the coordination structure. This question still needs to be further investigated in future research.

In this section, studies on the prosody of embedded clauses in Japanese have been reviewed. All in all, none of the data discussed here presents positive evidence for the effect of MatchClause. In the next section, we turn to a claim by Selkirk (2009), which specifically argues for the effect of MatchClause on embedded clauses.

3.2 *Wh*-prosody in Fukuoka Japanese

Selkirk (2009) claims, based on the literature on the prosody of *wh*-questions in Fukuoka Japanese (FJ hereafter), that this dialect of Japanese shows the mapping of standard clauses to ι and hence supports the syntax–prosody mapping at the clause-level. It will be shown in this section that there are a few empirical problems to this claim. Additional data from the relevant literature show the apparent clause–ι correspondence cannot be maintained as a general pattern of the *wh*-prosody in FJ. Crucially, the lack of correspondence is not due to some prosodic wellformedness constraint that suppresses the effect of MatchClause. The domain of the special *wh*-prosody (be it the one in FJ or that of Tokyo Japanese) is determined by two factors that are independent of MatchClause. The discussion leads to the conclusion that Japanese *wh*-question data do not constitute supporting evidence for MatchClause.

3.2.1 *Wh*-prosody in Tokyo Japanese

The prosody of *wh*-questions in Japanese has been discussed extensively since the early 2000's. Deguchi and Kitagawa (2002) and Ishihara (2002, 2003) claimed that in Tokyo Japanese (TJ hereafter), the scope of *wh*-question is marked prosodically by obligatory focus prosody. This obligatory *wh*-prosody starts from the *wh*-phrase and continues until the question particle[9] that binds the *wh*-phrase.[10]

[9] In TJ, the question particle is *ka*. While it is obligatory in the embedded clause (i.e., in an indirect *wh*-question), it may be omitted (i.e., phonetically null) in the matrix clause. The nominalizer *no* can, when it appears sentence-finally, be regarded as a question particle in the matrix clause (as in (19)). Cf. footnote 13 for question particles in FJ.

[10] See also Maekawa (1991, 1997) for earlier observations on *wh*-prosody in TJ.

This special *wh*-prosody in TJ is illustrated in Figure 2, which shows sample pitch contours for the sentences in (9).[11] The *wh*-phrase *dare* 'who' in the indirect *wh*-question (9b) shows a higher F_0-peak (= focal F_0-rise) compared to a non-*wh*-counterpart in (9a) *Naoya*, while the pitch contour of the post-*wh*-area *erimaki-o anda* 'knitted a scarf' is compressed (= post-focal reduction) until the question particle *ka* at the end of the embedded clause. The matrix material after the embedded clause *Yuuko-ni morasita* 'divulged to Yuko' shows the pitch range comparable to the one preceding the focus prosody, i.e., *Mari-wa Yumi-ga* 'Mari-TOP Yumi-NOM'.

(9) a. Indirect Yes/No-question

Mari-wa [Yumi-ga Naoya-ni erimaki-o anda ka] Yuuko-ni morasita
M.-TOP Y.-NOM N.-DAT scarf-ACC knitted Q Y.-DAT divulged
'Mari divulged to Yuko whether Yumi knitted a scarf for Naoya.'

b. Indirect *wh*-question

Mari-wa [Yumi-ga ↑**dare**-ni↓ erimaki-o anda ka] Yuuko-ni morasita
M.-TOP Y.-NOM who-DAT scarf-ACC knitted Q Y.-DAT divulged
'Mari divulged to Yuko who Yumi knitted a scarf for.'

In the *wh*-scope marking analysis by Deguchi and Kitagawa (2002) and Ishihara (2002, 2003), the domain of *wh*-prosody is determined by two factors. The first factor is the semantic scope of the *wh*-question, which is determined by the syntactic location of the question particle that semantically binds the *wh*-phrase. In the case of (9b), the scope of the indirect *wh*-question is the embedded clause, and hence, the *wh*-prosody ends at the end of the embedded clause. The second factor is the (linear) location of the *wh*-phrase within that scope. When the *wh*-phrase is at the clause-initial position, the *wh*-prosody starts at the beginning of the clause. When the *wh*-phrase appears clause-medially, as in (9b), the *wh*-prosody also starts sentence-medially, as in Figure 2.

As one of the crucial cases that tests the prediction of the *wh*-scope marking analysis, the prosody of scopally ambiguous *wh*-question sentences like (10) has been discussed and experimentally examined in the literature. In (10), the *wh*-phrase (located inside the embedded clause) can possibly be interpreted either at the embedded clause or at the matrix clause. In the former case, the sentence is a matrix Yes/No-question containing an indirect *wh*-question (i.e., the embedded *wh*-question reading, (10a)). In the latter case, it is a matrix *wh*-question containing an embedded Yes/No-question (i.e., the matrix *wh*-question reading, (10b)).[12]

[11] In the following TJ examples, the focal F_0-rise is indicated by ↑ at the beginning of a *wh*-phrase, and the post-focal reduction is indicated with ↓ at the end of the *wh*-phrase and the underline on the following words. The sentence-final rising contour, which is obligatory for matrix *wh*-questions in TJ as well as FJ, is not indicated, as it is not directly relevant for the discussion here.

[12] The English translation of (10b) violates the so-called *wh*-island condition. While the acceptability of (10b) varies among native speakers of Japanese and there is a general preference in acceptability toward the embedded *wh*-question reading in (10a), this reading is not considered ungrammatical in Japanese. See Kitagawa and Fodor (2003, 2006); Hirose

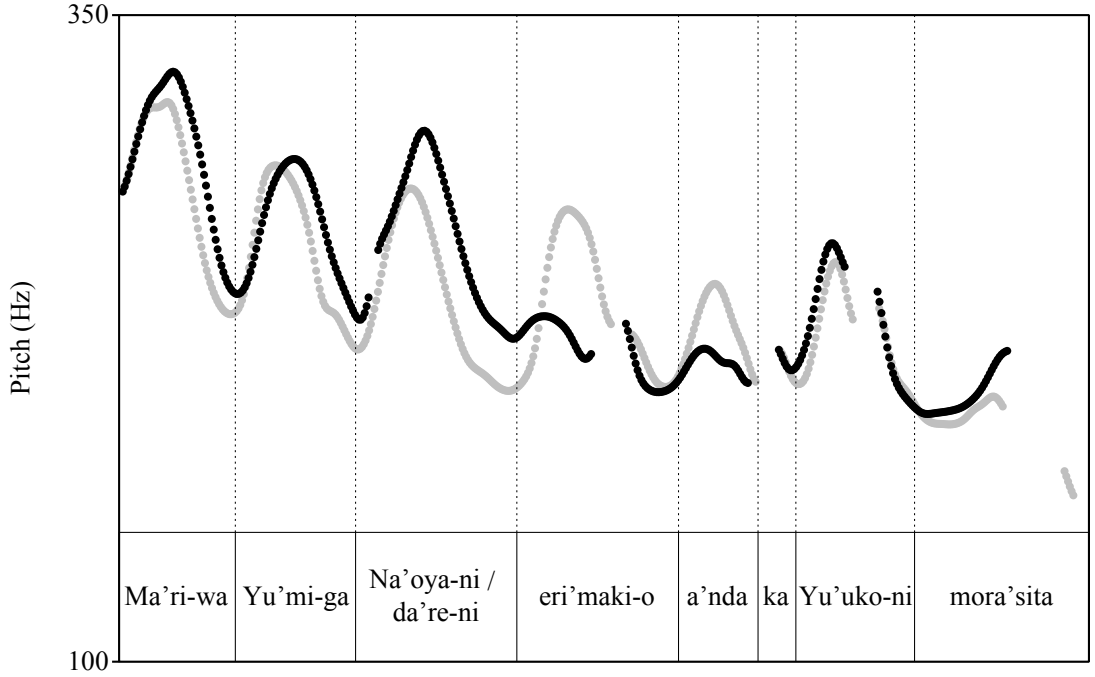


Figure 2: A sample pitch contour of (9b) (black line) superimposed onto that of (9a) (gray line).

(10) Scopally ambiguous *wh*-question sentence and the predicted *wh*-prosody

[Naoya-wa [Mari-ga **nani-o** nomiya-de nonda **ka**] Yumi-ni morasita **no**?]
N.-TOP M.-NOM what-ACC bar-LOC drank Q Y.-DAT divulge Q

a. Embedded *wh*-question:

'Did Naoya divulge to Yumi what Mari drank at the bar?'

[. . . [. . . ↑nani-o↓ . . . ka] . . . no?]

b. Matrix *wh*-question:

'(lit.) What$_i$ did Naoya divulge to Yumi whether Mari drank t_i at the bar?'

[. . . [. . . ↑nani-o↓ . . . ka] . . . no?]

Deguchi and Kitagawa (2002) and Ishihara (2002, 2003) claimed, based on their impressionistic observations, that the *wh*-prosody starts from the *wh*-word and ends at the end of the embedded clause when the sentence is intended as a matrix Yes/No-question containing an indirect *wh*-question (i.e., the embedded *wh*-question reading) as in (10a), while the *wh*-prosody continues until the end of the matrix clause when the sentence is intended as a matrix *wh*-question containing an embedded Yes/No-question (i.e., the matrix *wh*-question reading) as in (10b). That is to say, the scope ambiguity of sen-

and Kitagawa (2011); Hwang (2011); Kitagawa and Hirose (2012) for related discussion.

tences like (10) is disambiguated by the domain of *wh*-prosody. Most of the follow-up studies testing this claim have replicated the predicted prosodic patterns illustrated in (10a) and (10b) in their production experiments (Hirose and Kitagawa 2011; Hwang 2011; Kitagawa and Hirose 2012, with exception of Hirotani 2005, see below for related discussion).

3.2.2 Selkirk's (2009) analysis of *wh*-prosody in Fukuoka Japanese

FJ is also known to exhibit special prosody between the *wh*-phrase and the question particle[13] (Hayata 1985; Kubo 1989, 2001, 2005; Smith 2005, 2013, 2016; Hwang 2011). What is special about FJ is that the *wh*-prosody in this dialect is, unlike that of TJ, distinct from focus prosody and unique to *wh*-questions (Hwang 2011; Smith 2016). It is especially noteworthy that it overrides lexical pitch accent patterns by deleting any accentual F_0-falls between the *wh*-phrase and the question particle. The result is a high plateau starting from the *wh*-phrase (to be more precise, from the second mora of the *wh*-phrase, following the so-called *initial lowering*, Pierrehumbert and Beckman 1988) and ending right before the question particle (which may be phonologically null in matrix *wh*-questions in both FJ and TJ, see footnotes 9 and 13).[14]

(11) Declarative sentence in FJ (adapted from Kubo 2005:199; Selkirk 2009:56)

kyo⌝o bi⌝iru no⌝nda
today beer drank
'I drank beer today.'

(12) *Wh*-prosody in FJ (matrix) *wh*-questions (adapted from Kubo 2005:199; Selkirk 2009:56)

a. kyo⌝o dare-ga biiru nonda ∅
today who-NOM beer drank Q
'Who drank beer today?'

b. dare-ga kyoo biiru nonda ∅
who-NOM today beer drank Q
'Who drank beer today?'

(13) *Wh*-prosody in FJ indirect *wh*-question (adapted from Kubo 2005:201; Selkirk 2009:56)

[[[dare-ga kyoo biiru nonda] ka] sitto⌝o ∅]
who-NOM today beer drank Q know Q
'Do you know who drank beer today?'

Making the assumption that MatchClause applies to any clause in FJ (embedded or not), Selkirk

[13] In FJ, the question particle in the embedded clause is *ka*, as in TJ. The question particle for the matrix clause is either null (as in TJ, indicated here as ∅) or *to* (which is identical to the complementizer 'that' in TJ). Cf. footnote 9.

[14] In the following FJ examples, the location of F_0-falls (due to lexical H*+L pitch accents) is, whenever relevant, indicated with the symbol ⌝, and the high plateau of *wh*-prosody with overlines. The sentence-final rising contour is not indicated, as in the TJ examples (see footnote 11).

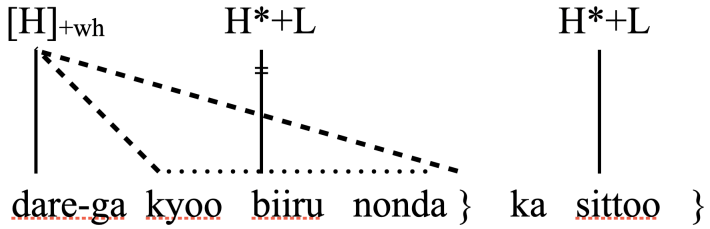
(2009:57) suggests the following ι-phrasing given in (14) and (15) for (12a) and (13), respectively:

(14) {$_\iota$ kyoo dare-ga biiru nonda ∅ }
who-NOM Q

(15) {$_\iota$ {$_\iota$ dare-ga kyoo biiru nonda } ka sittoo ∅ }
who-NOM Q Q

Selkirk's (2009) analysis is that *wh*-phrases carry a high tone *wh*-morpheme $[H]_{+wh}$ which has to be autosegmentally associated with the right edge of the ι containing the *wh*-phrase. The association of $[H]_{+wh}$ and the right edge of the ι causes the deletion of intervening lexical tones, yielding a high plateau, as illustrated in (16). In this analysis, it is predicted that the high plateau always ends at the right edge of the ι that immediately contains the *wh*-phrase.

(16) Deletion of lexical tones by multiple linking/spreading of $[H]_{+wh}$ morpheme (adapted from Selkirk 2009:58)



This generalization is empirically not correct, however. When there is a *wh*-phrase in an embedded clause, and the scope of *wh*-question is the matrix clause, as in (17) and (18), the high plateau does not terminate at the end of the embedded clause, where the right edge of an ι is predicted by MATCH-CLAUSE. Instead, the high plateau continues until the end of the matrix clause. The *wh*-prosody continues past the relative clause boundary in (17), and past the adjunct clause boundary in (18):

(17) [$_{CP}$ donna sigoto siyoo] hito-to ano-hito kekkonsita to?
what.kind job do person-with that-person got.married Q
'(lit.) The person [who does what kind of job] did that person get married to?'

'What kind of job is x such that that person got married to the person who does x?'

(Kubo 1989:81, glosses and translation added by S.I.)

(18) Ta⌝roo-wa [$_{CP}$ dare-ga kita tokorode] kaeru ∅?
T.-TOP who-NOM came at.which.point go.home Q
'Who is x such that Taro will go home when x comes?'

(Kubo 2001:29, glosses and translation added by S.I.)

If we assume that all embedded clauses are mapped to an ι and that the $[\text{H}]_{\text{wh}}$ gets associated at the right edge of that ι, it would wrongly be predicted that the high plateau triggered by the *wh*-phrase stops at the end of the embedded clauses in (17) and (18). Instead, together with the examples of indirect *wh*-questions like (13), where the *wh*-prosody terminates at the end of the embedded clause, these examples indicate that the *wh*-prosody of FJ marks the scope of *wh*-question, exactly like that of TJ.

Selkirk (2009:61) states that her ι-based account of the FJ *wh*-prosody cannot be extended to the *wh*-prosody in Tokyo Japanese, because it has been claimed in the literature that the post-*wh* pitch compression goes beyond the (embedded) clause boundary if the *wh*-phrase is in an embedded clause and the scope of *wh*-question is the matrix clause, as shown in (19):

(19) Matrix scope question with embedded *wh*-phrase

Naoya-wa [[Mari-ga ↑**nani**-o↓ nomiya-de nonda] to] Yumi-ni morasita no?
N.-TOP M.-NOM what-ACC bar-LOC drank that Y.-DAT divulged Q
'What did Naoya divulge to Yumi that Mari drank at the bar?'

The FJ examples in (17) and (18), however, show that FJ behaves exactly in the same way when the *wh*-phrase is in an embedded clause of a matrix *wh*-question.

In fact, Selkirk (2009:61) casts doubt on the *wh*-scope marking analysis (Deguchi and Kitagawa 2002; Ishihara 2002, 2003), based on the counterarguments made by Hirotani (2005). Hirotani examines the validity of the *wh*-scope marking analysis by conducting a series of production and perception experiments testing scopally ambiguous *wh*-question sentences like (10) in §3.2.1. One of the findings in her comprehension studies is that sentences like (10) with the matrix scope prosody (*wh*-prosody stretching until the end of the matrix clause) were interpreted as the matrix *wh*-question less than 50% of the time, while sentences with the embedded scope prosody (*wh*-prosody ending at the end of the embedded clause) were consistently (over 80%) interpreted as embedded *wh*-question. Based on these results as well as those from other experiments, Hirotani concluded that the predictions of the *wh*-scope marking analysis by Deguchi and Kitagawa (2002) and Ishihara (2002, 2003) were not borne out.

Hirotani's arguments against the *wh*-scope marking analysis, however, have been criticized by other researchers (see, in particular, Ishihara 2003:63–65; Hwang 2011:34–40, Kitagawa and Hirose 2012:633–635, for detailed discussion), due to several problems with the methodology of the experiments as well as her interpretation of the results. One of the problems pointed out by these researchers is that Hirotani disregards the overall preference toward the embedded scope reading in scopally ambiguous *wh*-question sentences like (10). For example, as pointed out by Hwang (2011:37), the unexpectedly low percentage of the matrix scope reading with the matrix *wh*-scope prosody seems (at least partly) due

to the fact that Hirotani's data includes the results from those speakers who did not accept the matrix readings at all.

The overall preference for the embedded *wh*-scope has been discussed in detail and experimentally confirmed in the literature (Kitagawa and Fodor 2003, 2006; Hirose and Kitagawa 2011; Kitagawa and Hirose 2012). In relation to this bias, studies also report asymmetries between production and comprehension of the *wh*-scope (i.e., asymmetry between speakers and listeners). For example, Kitagawa and Hirose (2012) have found that even though speakers mainly use *wh*-prosody to disambiguate the scope of the question, listeners rely more on the F_0-peak of the *wh*-phrase to distinguish the two scope readings. Kitagawa and Hirose (2012) argue that this asymmetry is related to the general preference for the embedded *wh*-scope reading in (10a) over the matrix *wh*-question reading in (10b). Once this bias is taken into consideration, Hirotani's results do not necessarily contradict the *wh*-scope marking analysis. Considering that the production data in other follow-up studies on this issue (Hirose and Kitagawa 2011; Hwang 2011; Kitagawa and Hirose 2012) are all in support of the *wh*-scope marking analysis, it seems reasonable to conclude that the *wh*-prosody in TJ functions as a *wh*-scope marker, contrary to the claim made by Hirotani (2005).

Furthermore, and more importantly to the current discussion, the *wh*-scope marking property has also been attested for FJ *wh*-questions in a production experiment conducted by Hwang (2011). In her production experiment, scopally ambiguous sentences parallel to (10) were distinguished prosodically, depending on the interpretation, as illustrated in Figure 3.[15]

In Figure 3, the embedded *wh*-question (solid line) shows a high plateau (indicated by the shorter $\leftrightarrow$), which is followed by a sharp F_0-fall (marked with a circle) at the embedded clause complementizer, i.e., the question particle. In the matrix *wh*-question reading (dashed line), by contrast, the high plateau (indicated by the longer $\leftrightarrow$) continues until the end of the sentence, followed by a H% boundary tone indicating the matrix question.

The prosodic pattern like the latter is not expected under Selkirk's (2009) analysis of FJ. Since MatchClause requires an ι-boundary at the end of the embedded clause, Selkirk's analysis would wrongly predict that the $[H]_{wh}$ on the *wh*-phrase would be associated with this ι-boundary. The resulting prosodic phrasing would be like the one for the embedded *wh*-question, not the one for the matrix *wh*-question, contrary to fact.

In summary, given the data from the literature on FJ, such as (17) and (18) above, as well as ex-

[15] In addition to the high plateau pattern exemplified in Figure 3, the speakers of Hwang's (2011) experiment, all college students, displayed other patterns that are not expected from the description of FJ *wh*-prosody in the literature. One speaker did not produce a high plateau but produced the TJ-type *wh*-prosody. Other speakers used unaccented *wh*-phrases as expected, but did not show the expected accent deletion (i.e., the high plateau ends at an intervening lexical accent). Hwang suspects that the prosodic system of FJ is changing, presumably due to the influence of TJ intonation. It should be added, however, that within the data with the expected high plateau pattern, the prosodic difference between embedded and matrix *wh*-questions shown in Figure 3 was statistically significant.

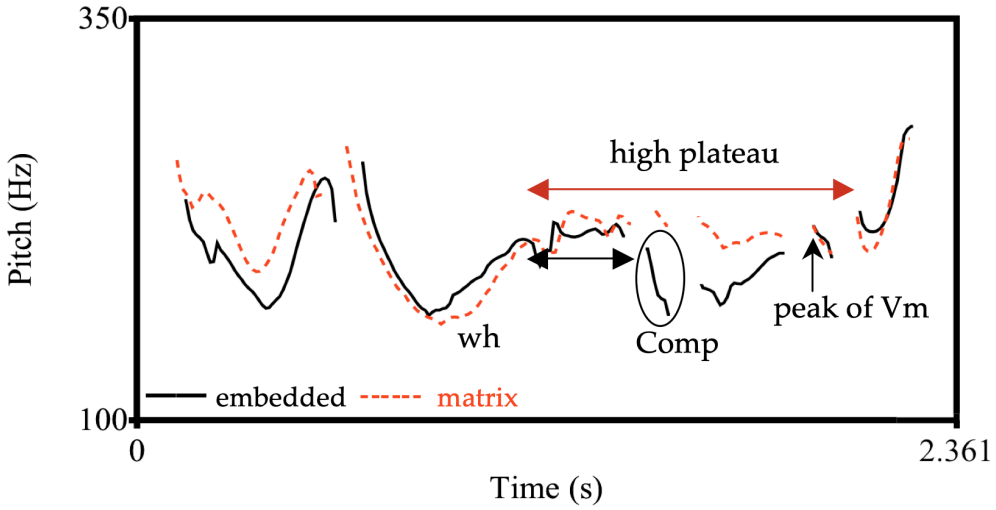


Figure 3: Sample pitch contours of a scopally ambiguous *wh*-question sentence in FJ that is parallel to (10), with the embedded *wh*-question reading (solid line) and with the the matrix *wh*-question reading (dashed line), taken from Hwang (2011:125). "wh," "Comp," and "Vm" denote *wh*-word, (embedded clause) complementizer, and the matrix verb, respectively.

perimental evidence from Hwang (2011), it is safe to conclude that the domain of *wh*-prosody is determined by the scope of the *wh*-question, as in TJ.

3.2.3 *Wh*-prosody in indeterminate constructions

In addition to the data we looked at so far, there is yet another set of facts that goes against Selkirk's (2009) claim that FJ provides evidence for the effect of MatchClause on standard clauses. The special *wh*-prosody that we have been discussing can be observed not only in *wh*-questions, but also in another *wh*-construction, namely, the so-called *indeterminate constructions* (Kuroda 1965, 2013; Nishigauchi 1990; Kishimoto 2001; Shimoyama 2001, among others). In this construction, *wh*-prosody appears between a *wh*-element and the quantificational particle *mo*.

Starting with TJ, there are two points worth mentioning regarding the *wh*-prosody found in indeterminate constructions in this dialect. First, unlike the case of *wh*-questions, where the *wh*-scope is always realized with focus prosody (i.e., focal F_0-rise on the *wh*-phrase, followed by post-focal reduction), there are two prosodic patterns that can appear between a *wh*-phrase and *mo* (Ishihara 2003:73–75, Kuroda 2013). Examples of the two patterns are shown in Figure 4, taken from Ishihara (2003:75), which are sample pitch contours for the indeterminate construction sentence in (20), produced by a single speaker. The first pattern is the focus prosody, just like in *wh*-questions of this dialect, as in (20a). The other pattern is, interestingly, a high plateau, identical to the *wh*-prosody in FJ *wh*-questions, as in (20b). The *wh*-prosody is marked with two different types of shades in Figure 4. In either pattern, the *wh*-prosody starts at the *wh*-phrase, and ends at the particle *mo*.

(20) aru nyuusu-wa [Nomo-ga **dare-ni** nakkuru-o nageta to-**mo**] ookiku
certain news-TOP N.-NOM who-DAT knuckleball-ACC pitched that-MO widely
hoozi-nakat-ta
broadcast-NEG-PST
'One news program did not widely broadcast that for any x, x is a person, Nomo pitched a knuckleball to x.'

a. *Wh*-scope indicated with focus prosody (= Figure 4 top)
a⌝ru nyu⌝usu-wa [No⌝mo-ga ↑**da⌝re-ni**↓ na⌝kkuru-o na⌝geta to-**mo**] o⌝okiku hoozi-na⌝kat-ta.

b. *Wh*-scope indicated with a high plateau (= Figure 4 bottom)
a⌝ru nyu⌝usu-wa [No⌝mo-ga **dare-ni** nakkuru-o nageta to-**mo**] o⌝okiku hoozi-na⌝kat-ta.

The second point worth mentioning, which is particularly relevant to the discussion here, is that unlike *wh*-questions, in which the end of *wh*-prosody always coincides with the end of a clause (where the question particle appears), an indeterminate construction can be formed at the level of the phrase as well as the clause. This is because *mo* can attach not only to a CP, as in (20), but also to a phrase (DP or VP), as in (21).

(21) Indeterminate constructions with quantificational particle *mo* attached to DP and VP (adapted from Kuroda 2013:67 and Kishimoto 2001:599, respectively, with emphases, pitch accents, and the high-plateau-type *wh*-prosody added by S.I.)

a. Ha⌝nako-wa [$_{DP}$ **dare**-ga hometa hon]-**mo** yom-a⌝nai
H.-TOP anyone-NOM praise book MO read-NEG
'Of anyone, Hanako does not read any book they praised.'

b. Ta⌝roo-wa [$_{VP}$ **nani**-o tabe]-**mo** si-na⌝kat-ta.
T.-TOP anything-ACC eat MO do-NEG-PST
'Taro did not eat anything.'

In both (21a) and (21b), *wh*-prosody is observed. And crucially, in both cases, *wh*-prosody starts from the *wh*-phrase, and ends with the particle *mo*, regardless of whether the particle attaches to a clause or a phrase.[16] When the high-plateau-type *wh*-prosody (like the one in FJ *wh*-questions) appears, the lexical accent of the *wh*-phrase *da⌝re/na⌝ni* and the following words (e.g., the verb *ta⌝be* in (21b)) will be deleted and a high plateau is formed between the *wh*-phrase and *mo*.

In the case of FJ, the data from the literature discussed above seem to show that indeterminate

[16]There is variation as to exactly where the plateau ends, because the domain of *wh*-prosody seems to be affected by various factors, such as length of the domain, the presence of intervening pitch accents, and the type of boundary pitch movement (BPM, see Igarashi 2015, 2018 and references therein) at the end of the domain. A closer examination is necessary to make a precise description. The variability of the domain, however, does not undermine the argument here, as the domain can be extended up to *mo*, but never beyond *mo*.

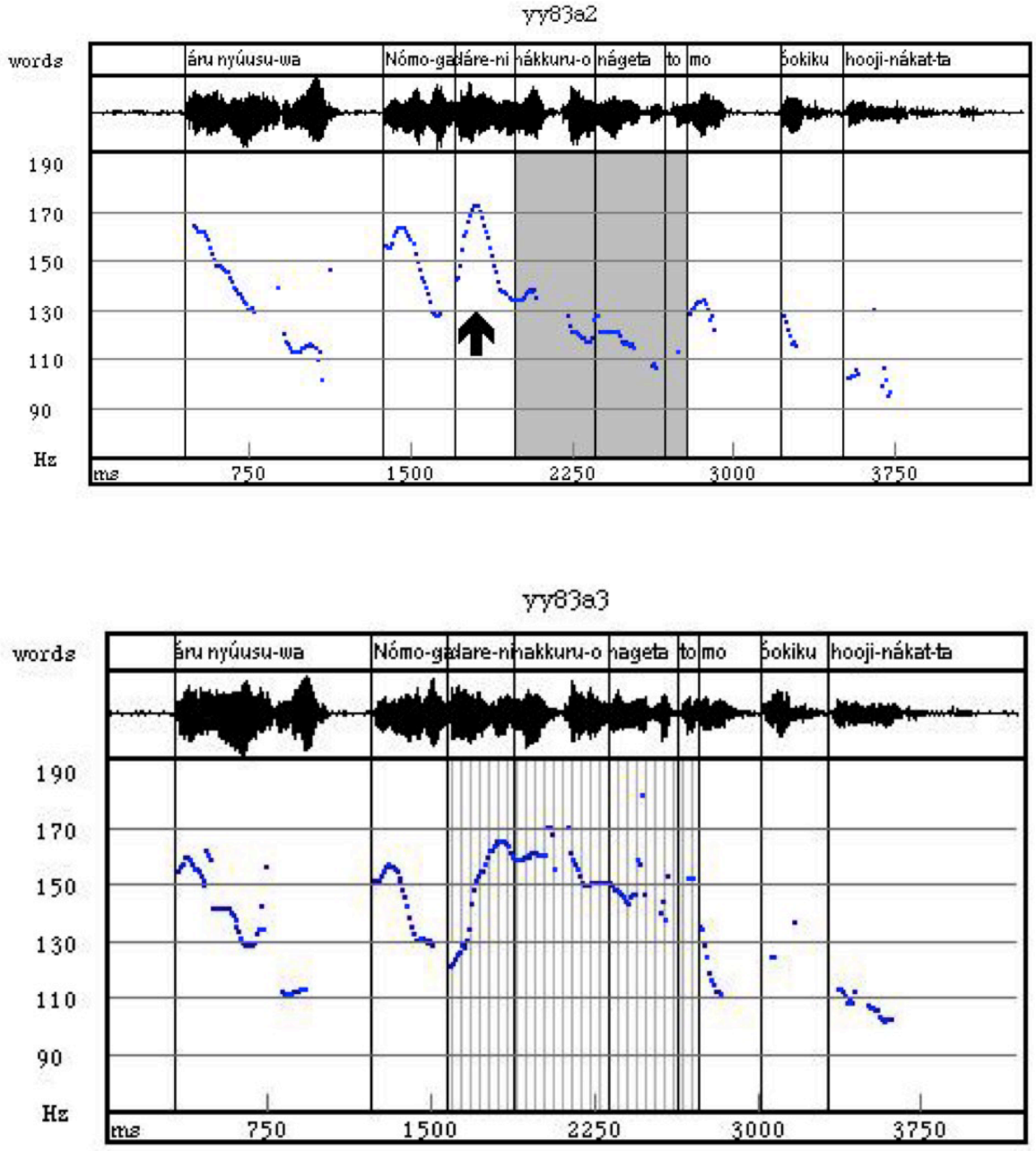


Figure 4: Two variants of *wh*-prosody in an TJ indeterminate sentence (20), produced by the same speaker, taken from Ishihara (2003:75). One is parallel to the *wh*-prosody in TJ *wh*-questions (top), while the other is parallel to the *wh*-prosody in FJ *wh*-questions (bottom).

constructions show the same *wh*-prosody as *wh*-questions. The following examples are from Kubo (1989:77–78, glosses and translations added by S.I.):

(22) a. [$_{CP}$ **itu** kyooto itte-**mo**] i'i yo
when Kyoto go-MO good PCL
'Kyoto is great whenever you visit there.'

b. [$_{DP}$ **dono** ziki-no kyooto-**mo**] i'i tokoro'-ga aro'o ga
which season-GEN Kyoto-MO good place-NOM exit PCL
'Whatever the season in Kyoto, there are great places [to visit].'

Cases like (22b), where the domain of *wh*-prosody is a DP rather than a clause, cannot be accounted for by Selkirk's (2009) analysis of *wh*-prosody, because the domain of *wh*-prosody does not correspond to the domain predicted by MatchClause. The domain of *wh*-prosody in (22b) is the DP *dono ziki-no kyooto-mo*, while MatchClause only predicts that there is an ι that corresponds to the entire sentence. MatchPhrase does not predict the presence of an ι-boundary at the end of the DP, either, as it only predicts the presence of φ-boundaries.

The data from indeterminate constructions further strengthens the *wh*-scope marking analysis of *wh*-prosody in Japanese, for both TJ and FJ. The domain of *wh*-prosody is determined by the syntactic location of the relevant particle, which defines the scope of a question/quantification. Because the scope domain may be a phrase or a clause (in the case of indeterminate constructions), the analysis based on MatchClause cannot account for the data.

In sum, we reviewed Selkirk's (2009) claim that the FJ *wh*-prosody can be explained as an effect of MatchClause on embedded clauses. A further look at the data in literature shows that her observation is empirically not well grounded. The *wh*-prosody in FJ, be it the one in *wh*-questions or the one in indeterminate constructions, behaves like the *wh*-prosody in TJ in terms of how its domain is determined: it functions as a *wh*-scope marker rather than being formed as a result of MatchClause.

3.3 Possible approaches to (apparent) ι-mapping of embedded clauses

So far, we have seen that the Japanese data in previous studies do not constitute empirical support for Match(clause, ι). As correctly pointed out by an anonymous reviewer, however, there are several studies that have presented evidence for the presence of the mapping of embedded clauses to ι's in other languages. For example, there are studies that clearly indicate the presence of ι-boundaries at embedded clause boundaries, such as Truckenbrodt (2005), Truckenbrodt and Darcy (2010) and Schubö (2020), for German. Schubö (2020) proposes a Match-based account of the clause-level mapping observed in his data. In the remainder of this section, I briefly comment on cases that are consid-

ered to be evidence for the mapping of embedded clauses onto ι's.

Cross-linguistically, empirical evidence for the clause-level mapping is, at least at this point, much more scarce in comparison to the phrase-level mapping. One possible reason is that research on clause-level intonation is still at an early stage in comparison to word- and phrase-level intonation. At the same time, it is also well-known by now that embedded clauses often fail to map to ι in many languages (including Japanese, as shown above). There are also cases where the mapping of embedded clauses to ι's is found to be optional in a language, as in the case of Stockholm Swedish (Myrberg 2010, 2013).

If the mapping of embedded clauses to ι is only partial, there seem two possible approaches to choose between. One is to assume that there is a general constraint on the clause-level syntax–prosody mapping, as proposed in Match Theory, and to try to account for the occasional lack of clause-level mapping. Another possible approach would be to assume that there is no such constraint as MATCH-CLAUSE, and to try to account for the cases of (apparent) clause-level mapping using other constraints.[17] As will be argued in section 4, MATCH(illocutionary clause, ι) should be replaced by another constraint. Then, it is plausible to hypothesize that apparent match effects found in embedded clauses are derived by other constraints. When testing MATCHCLAUSE empirically, it is therefore important to take into consideration not only those factors that may suppress the effect of MATCHCLAUSE, but also factors that may derive the same effect as MATCHCLAUSE even in the absence of MATCHCLAUSE.

Even when a consistent correspondence between embedded clauses and ι's is found in a dataset, it is important to examine whether it is the clausehood that is responsible for the observed correspondence, or some other factors that derive it. For example, in the case of relative clauses, it should be noted that relative clauses are always embedded in an NP headed by the head noun of the relative clause. This means that one of the syntactic edges (either the left or the right one, depending on the language) of a relative clause coincides with the edge of the NP containing it. Because of that, any syntax–prosody correspondence found at this position may be due to the clause boundary of the relative clause, or the phrase boundary of the NP containing the relative clause, or both. This means that, even when a relative clause shows a correspondence to an ι in some language, it is not clear whether the correspondence is between the relative clause itself and the ι, or between an NP that contains the relative clause and ι (unless a prosodic boundary also appears on the other end, i.e., between the head noun and the relative clause).

Another point, which applies to any kind of embedded clause, that should be taken into consideration is the length, as well as the internal structure, of relevant constituents. When looking at previous

[17] The same point was made by Hamlaoui and Szendrői (2017), who assume a mapping constraint that only refers to root clauses, and do not assume any mapping constraint for non-root clauses.

studies, it is sometimes not entirely clear whether it is the difference in the syntactic category (phrase vs. clause) or the general length (short vs. long constituent) that is affecting the prosodic phrasing.

Truckenbrodt (2005), for example, convincingly shows that the right edge of an embedded clause in German consistently coincides with the right edge of an ι. He examined the prosody of sentences containing a subject clause, a (restrictive) relative clause, or an object clause, in different syntactic positions. He also compared root and embedded coordinated clauses. In all of these cases, the right boundary of an ι is found consistently at the right edge of a clause, be it a matrix one or an embedded one.

It should be noted, however, that the embedded clauses contained in the test sentences make the relevant constituent relatively long, as each clause contains at least two or more content words that bear a pitch accent, which means that the prosodic constituent containing (or corresponding to) the relative clause dominates at least two φ's. (23) is an example with a relative clause in a sentence-medial position:[18]

(23) [$_{\text{CP}}$ Die Lola hat dem Mann, [$_{\text{CP}}$ der einem Maurer einen Löwen gemalt hat],
the.NOM L. has the.DAT man REL.NOM a.DAT mason a.ACC lion painted has
ein Lob gegeben]
a.ACC praise given
'Lola praised the man who painted a lion for a mason.'

(Truckenbrodt 2005, punctuation marks, glosses and translation added by S.I.)

If the ι-boundary found at the end of the relative clause is the result of MatchClause (or a comparable Alignment constraint, as proposed by Truckenbrodt 2005), it would be predicted that there will be no ι-boundary if the relative clause is replaced by XPs of the same length. If the same kind of ι-boundary is found, contrary to the prediction, they are likely to be due to some other factor.

It might be possible to hypothesize that an ι-boundary may be indicating the start and/or the end of a relatively long XP, as a result of prosodic promotion of a long φ that dominates many φ's to an ι. Even though prosodic recursion (e.g., a φ dominating another φ) is generally allowed in the prosodic structure assumed in Match Theory as well as in Prosodic Adjunction Theory, there may be some kind of prosodic wellformedness constraint on the number of daughter constituents of the same category. If such a constraint is assumed, a φ dominating a small number of φ's as in (24a) would be a more optimal structure than a φ dominating more φ's as in (24b). If, however, the φ containing a large number of φ's is promoted to ι, as in (24c), it avoids the violation of such a constraint.

(24) a. $(_{\varphi}\ (_{\varphi}\ \ldots\ )\ \ldots\ \ldots\ )$ — OK

[18] Besides the length, it might potentially be relevant as well that commas at clause boundaries are obligatory in German (for both restrictive and non-restrictive relative clauses, as well as for clausal ojects).

b. $(_{\varphi}\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ )$ — Not OK

c. $\{_{\iota}\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ \}$ — OK

As a concrete example, we can consider German object clauses. If we assume a model that entirely lacks MatchClause, a German sentence with an (extraposed) object clause is predicted to be mapped to a prosodic structure like (25).[19] Here, I am assuming, for the sake of exposition, that Match(illocutionary clause, ι) is at work, mapping the ForceP onto an ι. See §4 for relevant discussion.

(25) Syntax: $[_{\text{ForceP}}\ [_{\text{CP}}$ main clause $]\ [_{\text{CP}}$ embedded clause $]\ ]$
Prosody: $\{_{\iota}\ (_{\varphi}\ )\ (_{\varphi}\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ )\ \}$

Depending on the prosodic structure of the clause-internal part, one of the clauses may be realized as an ι, due to the prosodic wellformedness constraint on the number of the daughter constituents suggested above. The resulting structure would be something like (26):

(26) Syntax: $[_{\text{ForceP}}\ [_{\text{CP}}$ main clause $]\ [_{\text{CP}}$ embedded clause $]\ ]$
Prosody: $\{_{\iota}\ (_{\varphi}\ )\ \{_{\iota}\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ \}\ \}$

Once one of the clauses is promoted to an ι, the other clause is likely to be promoted to ι as well, due to EqualSisters (Myrberg 2013), another prosodic wellformedness constraint defined as in (27):

(27) EqualSisters (Myrberg 2013:75)

Sister nodes in prosodic structure are instantiations of the same prosodic category.

The resulting prosodic structure will be as in (28) below:

(28) Syntax: $[_{\text{ForceP}}\ [_{\text{CP}}$ main clause $]\ [_{\text{CP}}$ embedded clause $]\ ]$
Prosody: $\{_{\iota}\ \{_{\iota}\ \}\ \{_{\iota}\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ (_{\varphi}\ \ldots\ )\ \}\ \}$

Another possible account of prosodic promotion of two sister φ's to ι's may be what Selkirk (2005) calls *stylistic promotion*. Selkirk (2005) suggests that, for stylistic purposes, φ's may sometimes be promoted to ι's. Stylistic promotion is suggested for cases where phrases in a single clause are parsed as ι's. According to Selkirk (2005:25–27), the sentence in (29) can be produced as a single ι containing two φ's, as in (29a) or as an ι containing two subordinate ι's, as in (29b):

(29) Three mathematicians in ten derive a lemma.

[19] In the case of German in particular, object clauses are considered to be extraposed, possibly to different landing sites, as suggested by Truckenbrodt and Darcy (2010). If this assumption is correct, then it is again unclear whether it is the syntactic status as a clause or as an extraposed constituent that triggers mapping to ι.

a. $\{_{\iota}$ ($_{\varphi}$ Three mathematicians in ten) ($_{\varphi}$ derive a lemma) $\}$

b. $\{_{\iota}$ $\{_{\iota}$ Three mathematicians in ten$\}$ $\{_{\iota}$ derive a lemma$\}$ $\}$

While the phrasing pattern in (29a) can be derived by the application of MatchClause and Match-Phrase, the phrasing pattern in (29b) cannot be derived by Match constraints alone, because the subordinate ι's correspond to phrases but not clauses. Selkirk proposes that the phrasing in (30b) is a result of stylistic promotion.

If such a mechanism exists as a general prosodic operation, it should in principle be applicable to φ's that correspond to (embedded) clauses as well. In the case of (25)–(28) above, the prosodic structure in (28) can be derived from the one in (25) as a result of stylistic promotion. Importantly, MatchClause is not involved in the process. Furthermore, the optionality of two phrasing patterns, as found in Stockholm Swedish, could also be explained under this analysis. Although details of the analyses suggested here still need to be elaborated, it seems possible to derive desired prosodic structures without resorting to MatchClause.

One prediction of the analyses without MatchClause is that the mapping of embedded clauses onto ι's is universally optional. Whether an embedded clause maps to an ι or not is determined purely in terms of prosodic wellformedness (or stylistic preference). Furthermore, the prosodic behavior of an embedded clause is, all other things being equal, predicted to be indistinguishable from that of a phrase. This contrasts with the prediction of Match Theory that there may be cases where clauses, but not phrases, map to ι. Once again, further research is needed to test this prediction.

4 Toward a theory of discourse–ι mapping

In the previous section, studies on the prosody of embedded clauses in Japanese, as well as FJ and TJ *wh*-prosody were examined. Based on the discussion, it was concluded that Japanese does not provide evidence for (the general version of) MatchClause. This suggests that the prosody of embedded clauses in Japanese can be explained without MatchClause (assuming that *wh*-prosody is regulated by other constraints). A possible grammatical model without MatchClause that can account for the apparent clause-level mapping in other languages was also discussed.

The main concern of this section is the other, more specific version of MatchClause, namely, Match(illocutionary clause, ι). The section starts with the discussion of the incompleteness as well as the inexhaustiveness of the ι-mapping predicted by MatchClause. The clause-level syntax–prosody mapping postulated via MatchClause is incomplete, in the sense that there are clauses that do not map to ι on a consistent basis, as we have already seen in §3. Besides that, it is also inexhaustive, in

the sense that there are cases where clauses and phrases both map to ι alike. This issue is related to the question regarding the relevance of illocutionary force, or speech act, to the subcategorization of clause types mentioned in §2.1. The discussion in this section leads to the idea that it is not the clause (as a syntactic category), but rather the speech act associated with a syntactic constituent (clause or phrase), that shows systematic correspondence to ι. It will be proposed, following recent proposals in the literature, that MATCH(illocutionary clause, ι) should be replaced by mapping of speech acts onto ι's.

4.1 Incompleteness and inexhaustiveness of ι-mapping

The question to be discussed here is whether it is really syntax that is related to ι-phrasing, or whether ι-mapping should instead be attributed to something else (that, in turn, may be associated with a syntactic constituent). Upon discussing this question, it is important to note that, in comparison to the well-established phrase-level correspondence (syntactic phrases $\rightarrow$ φ's), the clause-level correspondence is far less consistent (in the sense that not all clauses map to ι, as discussed in §3) and far less exhaustive (in the sense that non-clauses also map to ι).

In fact, finding a consistent mapping relation between syntactic constituents and ι has long been a challenge. Match Theory is not the only generalization that has been proposed in the literature. Selkirk (1984)'s *Sense Unit Condition* is one such attempt. Selkirk (2005) also attempts to characterize the mapping between syntactic Comma Phrase and ι, before shifting her theory toward Match Theory. The major reason for the difficulty seems to be the incompleteness and inexhaustiveness of the ι-mapping.

As for incompleteness, it has already been acknowledged by Selkirk (2011:452) that not all types of clauses show consistent correspondence to ι, as explained in §2.1. While what Selkirk (2011) calls illocutionary clauses seem to show a strong tendency to be mapped to ι in their prosodic representations, so-called standard clauses (mostly embedded clauses that do not have an illocutionary force of their own) do not show systematic correspondence. As already shown in §2.1, this itself can be explained under Match Theory, by assuming that some prosodic wellformedness conditions may suppress the effect of MATCH(standard clause, ι). As discussed in §3.3, however, the incompleteness of (embedded) clause–ι correspondence also raises the question of whether (the general version of) MATCHCLAUSE is needed in order to account for the syntax–prosody mapping of embedded clauses.

As for inexhaustiveness, there are cases where certain phrases show a strong tendency to be mapped to an ι. In such contexts, clauses and phrases map to ι alike. Two major areas can be mentioned in which ι-mapping is often discussed. The first area is parenthetical expressions, such as adverbial clauses/phrases,

non-restrictive relative clauses, appositive phrases, sentence adverbs, etc. (30) are some of the examples from Dehé (2014:3–6).[20] In particular, it should be noted that parenthetical expressions are not limited to clauses.

(30) a. When we were on holiday—*that reminds me, I must pick up the photos*—we saw so many interesting places. (Clauses)

b. A surprise present, *a bouquet of roses*, was delivered to my door. (Nominal apposition)

c. The secretary *well-mannered as anybody* will present an apology. (Lexical phrases, in this case, an AP)

d. Is it safe, *would you say*? (Interrogative parentheticals)

e. He suffered great mental distress *didn't he* after the war. (Question tags)

f. He is, unfortunately, ill. (Sentence adverb)

g. I've been dreaming of winning a gold medal for *what* 20 years now. (One-word expressions)

h. He is *oh!* so smart. (Interjections)

(Dehé 2014:3–6)

Among various types of parenthetical expressions, Dehé (2014) examined the prosodic phrasing of six types of parenthetical expressions (parenthetical clauses, non-restrictive relative clauses, nominal appositions, comment clauses, reporting verbs, and question tags), using the spontaneous and semi-spontaneous speech data from the British Component of the International Corpus of English (Nelson et al. 2002). The results show that the majority of the parenthetical clauses (99%), non-restrictive relative clauses (92%), nominal appositions (80%), and question tags (73%) are separated from the main (root) clause by ι-boundaries. The data suggest that certain clausal parentheticals (such as parenthetical clauses and non-restrictive relative clauses) and phrasal parentheticals (nominal appositive) behave alike.[21]

Another area in which ι has often been claimed to be involved is information-structural phenomena, most prominently, topicalization. Discourse topics tend to be prosodically separated from the rest of the sentence by a strong prosodic boundary, presumably an ι-boundary (among others, Frascarelli 2000; Feldhausen 2010).[22] Just like parenthetical expressions, topics may be a clause or a phrase. The

[20] See Dehé (2014) for an extensive overview of various types of parenthetical expressions, as well as the original sources of the examples in (30).

[21] Dehé (2014) also shows that not all parenthetical expressions are consistently mapped to ι. Comment clauses, for example, were separated by ι-boundaries from the main clause in only 23% of her data. They tend to be prosodically integrated instead. See below for relevant discussion.

[22] As an anonymous reviewer points out, there are different observations and claims with respect to the prosodic nature of the boundary found after topics. In some cases, post-topical prosodic boundaries are analyzed not as ι's but as intermediate phrase boundaries. As it is beyond the scope of this paper to examine each proposal in detail, I will leave this

syntactic difference between clauses and phrases does not seem to be playing any important role in mapping of these elements to an ι.

4.2 Connection between clauses and illocutionary force

As mentioned in §2, Selkirk (2011) proposes two types of MatchClause constraints. In this analysis, MatchClause (unlike MatchWord and MatchPhrase) is sensitive to the type of clauses (standard vs. illocutionary clause). What distinguishes between the two clause types is the presence or absence of illocutionary force. A clause that carries an illocutionary force of its own is categorized as an illocutionary clause. Due to the (universally high-ranked) Match(illocutionary clause,ι), it is (obligatorily) mapped to an ι. In contrast, a clause that does not carry an illocutionary force is categorized as a standard clause, and it may or may not be mapped to an ι, depending on the relative ranking of (the general version of) MatchClause and other prosodic wellformedness constraints.

Here, one may ask why clauses, but not phrases, need to be distinguished depending on whether they carry an illocutionary force or not. This question is highly relevant, because it seems that a parallel pattern can be found for phrases. As mentioned above, certain parenthetical expressions, which may be a clause or a phrase, are consistently mapped to an ι (See Dehé 2014 and references therein). Similarly, certain types of topics tend to be phrased as an ι, be it a clause or a phrase. What matters in these cases, then, seems to be the status as parenthetical expressions or as topics, rather than whether the expression is a clause or a phrase.

There seems to be one common trait among (certain) parenthetical expressions and (certain) topics: they constitute a speech act that is independent of that of the main (root) clause. It has often been claimed that parenthetical expressions express a separate speech act that is independent of that of the main clause. The illocutionary force of a parenthetical expression has also been observed to be independent of that of the main clause (see Dehé 2014 and references therein). Potts (2005) claims that parentheticals (or what he calls supplemental expressions) are syntactically marked as Comma Phrase and carry a conventional implicature, which, Selkirk (2005, 2009, 2011) assumes, is equivalent to saying that a phrase associated with a [+comma] feature performs a separate speech act.[23]

Interestingly, there have also been claims and suggestions that certain types of topic phrases perform a separate speech act. Ebert (2009), following Jacobs (1984), suggests that establishing a(n aboutness) topic is a separate illocutionary act. Bianchi and Frascarelli (2010) propose, adopting Krifka's (2001) analysis of quantified NPs, that aboutness-shift topics perform speech acts of their own.

issue for future research. It is enough at this point to note that there seems a general, cross-linguistic tendency to find a prosodic break after topic phrases, and that some of these cases are analyzed as ι-boundaries.

[23] Güneş (2014), however, points out problems of Potts's (2005) analysis of parenthetical expressions, as well as Selkirk's (2005; 2009; 2011) interpretation of Pott's Comma Phrase. See below.

Coming back to the formulation of MatchClause, if the presence or the absence of illocutionary force is the only difference between the two clause types (illocutionary and standard clauses), and their difference in terms of ι-mapping is explained solely based on this difference, it seems more reasonable to connect illocutionary force itself (or the speech act associated with it) and ι directly, rather than to attribute the ι-mapping to the clause type. It makes even more sense once we consider the fact that the phrasal expressions that tend to map to ι (such as parenthetical expressions and aboutness topics) are also claimed to perform separate speech acts.

In relation to this, Selkirk (2011:452) assumes that ForceP, the projection that dominates illocutionary clauses, is an instance of Potts's (2005) Comma Phrase. One problem is that, while it seems quite plausible to assume that ForceP and CommaP share the property of performing a speech act, their syntactic categories cannot be entirely equivalent. While ForceP is unambiguously a subtype of clauses, Comma Phrases subsume both clauses and phrases, as different types of parenthetical expressions are analyzed as instances of a Comma Phrase.[24]

Given that, instead of hypothesizing that there are two MatchClause constraints (and possibly along with other constraints that would additionally be needed to account for the cases of phrases that map to ι), it seems more reasonable to hypothesize that the illocutionary force, or the speech act associated with it, is the cause of ι-mapping. Once it is hypothesized that each speech act maps to ι, it would no longer be necessary to assume that clauses with illocutionary force map to ι, as proposed by Selkirk (2011).

In fact, such a proposal is not new. Güneş (2014) has already pointed out some of the observations made above. She investigated the prosodic phrasing of phrasal and clausal parentheticals in Turkish (epistemic adverbial *bence* 'for me' and epistemic conditional *yanılmıyorsam* 'if I am not wrong') and compared them with other kinds of parentheticals. The results of her investigation is summarized in Table 1.

	Type of parenthetical	Prosodic category
Phrasal	Vocative	ι
	Epistemic adverbial	φ
	Concessive phrase	φ
Clausal	Interruption (isolated)	ι
	Interruption (integrated)	ι
	Comment clause	φ
	Epistemic conditionals	φ

Table 1: Phrasing patterns of phrasal and clausal parentheticals in Turkish, taken from Güneş (2014:305).

Table 1 shows that both clausal and phrasal parentheticals map to ι. It also shows that not all par-

[24] This problem was also acknowledged by Güneş (2014).

entheticals map to ι. If all parentheticals share the same syntax—whether as a Comma Phrase (Potts 2005) or as a special kind of syntactic phrase, ParP, as proposed by de Vries (2012) (which is the analysis adopted by Dehé 2014 and Güneş 2014)—a mapping principle that maps all parenthetical constituents to ι, would make a wrong prediction for those that do not map to ι, e.g., comment clauses. Güneş (2014) instead claims that it is the presence or absence of speech acts that draws a line between parentheticals that map to ι and those that do not.

Truckenbrodt (2015) also shows, via tests using modal particles and sentence adverbs in German that induce a speech act distinct from that of the root clause, that various kinds of constituents (coordinated phrases and clauses, appositive relative clauses and nominal appositives, afterthoughts, peripheral adverbial phrases, parentheticals) that carry a speech act consistently map onto ι, be it a phrase or a clause. When these constituents do not carry a speech act, however, they do not have to be separate ι's. Truckenbrodt postulates the following descriptive generalization.

(31) Each speech act requires a separate intonation phrase and concomitant sentence stress.
(Truckenbrodt 2015:313)

Note that even though Truckenbrodt (2015) advocates Selkirk's (2011) core insight that speech acts show correspondence to ι's, his formulation in (31) does not refer to the clausehood of the constituent.

In the remainder of this section, this line of analysis will be explored, as well as its implications for the theory of syntax–prosody interface.

4.3 Mapping of speech act to ι

Although it is beyond the scope of this paper to develop a full-fledged analysis of the possibility introduced in the previous subsection, I will adopt, following Güneş (2014, 2015) and Truckenbrodt (2015), the following principle, for the sake of concreteness:

(32) A speech act is realized as an ι in the prosodic representation.

We may be able to formulate it as a MATCH constraint, MATCH(speech act, ι), or MATCHSA for short (see also Güneş 2014, 2015; Hamlaoui and Szendrői 2017).[25] This constraint covers all the cases of

[25] Güneş (2014, 2015) modifies Selkirk's (2011) MATCH (illocutionary clause, ι) by replacing Rizzian ForceP, which represents sentential force and is hence assumed to be constantly a clausal projection, with *illocutionary force projection* ($\text{Force}_{\text{ill}}$P), which heads all kinds of syntactic structures that are performed as speech acts and is hence not related to clausehood. The formulation in (i) can be considered a syntactically defined version of MATCHSA.

(i) $\text{MATCHFORCE}_{\text{ILL}}$ (Güneş 2014:283)
$\text{Force}_{\text{ILL}}$P in syntactic constituent structure must be matched by a constituent of a corresponding prosodic type, call it ι, in the phonological representation.

root clauses and parenthetical expressions (both phrasal and clausal ones), as well as discourse topics, in a uniform fashion. All root clauses are commonly assumed to have illocutionary force, and hence to perform a speech act. Parenthetical expressions and discourse topics that map to ι's, as discussed in §4.2, can also be analyzed as separate speech acts that are not part of the proposition expressed by the root clause. All these speech acts form separate ι's, according to MatchSA. Once MatchSA is assumed, it is no longer necessary to assume Match(illocutionary force, ι), as it is otherwise redundant.[26]

In essence, this proposal is closer to Selkirk's (2005) account, in which edges of Comma Phrases are aligned to edges of ι, than to Match Theory in Selkirk (2011), in the sense that ι-mapping is claimed to be related to discourse-related notions rather than the distinction between clause and phrase. In this regard, it is also interesting to note that Selkirk (2005:12–13) suggests that topic-like expressions, such as *as for*-phrases and left dislocated phrases, may count as expressions marked by [+comma] feature. As pointed out by Güneş (2014, 2015), however, the CommaP analysis of Potts (2005) adopted by Selkirk (2005) cannot distinguish parentheticals that perform a speech act (and map to ι) from those that do not. Some modification may be needed to successfully account for all the relevant cases.

Although there are problems in the analyses in Selkirk (2005) and Selkirk (2011), it is also important to note that they both relate ι-mapping to illocutionary force. Truckenbrodt (2015) also advocates Selkirk's (2011) Match account in this regard. Although they differ in details and core foci of the analyses, proposals that were discussed here, namely, Potts (2005), Selkirk (2005, 2011), Güneş (2014, 2015), Truckenbrodt (2015), Hamlaoui and Szendrői (2015, 2017) all seem to agree that the speech act plays a crucial role in ι-mapping. MatchSA captures this insight.

The conclusion that Match(illocutionary clause, ι) can be replaced by MatchSA is also related to the question whether the general version of MatchClause, which is responsible for the mapping of embedded clauses to ι's, should be maintained. Note that MatchSA, unlike MatchClause, is irrelevant for the cases where embedded clauses without illocutionary force (i.e., standard clauses) form an ι. Although Japanese does not show any evidence for MatchClause, as discussed in §3, as long as some languages exhibit mapping of embedded clause to ι, there remains a question whether Match-

Hamlaoui and Szendrői (2017) also assume a similar constraint (albeit in addition to the syntax–prosody mapping from root clause to ι). Hamlaoui and Szendrői's (2017) formulation, however, slightly differs from the one adopted here. Instead of exact correspondence (i.e., match) between speech acts and ι's, a speech act only needs to be contained in an ι. It remains to be seen which formulation yields more empirically correct predictions.

(ii) SpA-ι: Each Speech Act is contained in a single ι. (Hamlaoui and Szendrői 2017:5)

[26]The proposal in this paper differs from Hamlaoui and Szendrői (2017) in this regard. They claim that only root clauses, or to be more precise, the highest verbal projections (HVP) in their analysis, map to ι. This mapping is proposed in addition to the mapping of speech act to ι. Since root clauses are usually considered to perform speech acts, however, there seems no need to assume separate constraints for the mapping of speech acts and the mapping of root clauses, as long as root clauses are defined appropriately in terms of HVPs.

Clause should still be needed.

In §3.3, there was a discussion of how to account for cases where a constituent does not perform a speech act, but maps to ι. It was suggested that the (apparent) ι-mapping of embedded clauses may be explained without (the general version of) MatchClause. Although further investigation is needed to find out if such an approach is correct, it does seem a plausible hypothesis that is worth testing.

4.4 Prosodic hierarchy and division of labor

MatchSA is *not* a syntax–prosody mapping constraint, because speech acts are not a syntactic category like words and phrases.[27] If we assume that ι-mapping is regulated according to (32) (together with other prosodic wellformedness constraints), and that there is no MatchClause, it follows that the prosodic category ι is outside the domain of the syntax–prosody mapping.

In the standard theories of the syntax–prosody interface, Match Theory and Prosodic Adjunction Theory to be more precise, the prosodic hierarchy is divided into two categories, *rhythmic categories* and *interface categories*. Moras (μ), syllables (σ) and feet (f) belong to the former, while ω', φ's, and ι's belong to the latter. What the current proposal implies is that interface categories can further be divided into two (sub)categories. While the lower two prosodic categories, ω and φ, are responsible for the interface with syntax, the higher category, ι, is responsible for the interface with discourse/pragmatics. If this is the architecture of the interface between prosody and other modules of grammar, it seems natural that it has long been a challenge to come up with a convincing theory of the syntax–prosody mapping in relation to ι. A possible reason is that there are two different sources of ι's. One is the discourse–prosody mapping regulated by MatchSA, as proposed in §4. The other is prosodic promotion of φ to ι, caused by prosodic wellformedness (or stylistic promotion), as suggested in §3.3. Crucially, neither of these sources is (directly) related to the syntax–prosody mapping.

5 Concluding remarks and remaining questions

5.1 Summary

In this paper, the necessity and the validity of the two versions of MatchClause constraints were discussed, mostly in light of Japanese data. First, the mapping of embedded clauses onto ι's was discussed. It was shown that Japanese data from previous studies do not provide positive support for MatchClause. Relative clause data are inconclusive (Uyeno et al. 1979, 1981; Kubozono 1993; Hi-

[27] The proposal does not exclude, however, possibilities that speech acts are realized in the syntactic representation as a certain syntactic projection, such as $\text{Force}_{\text{ill}}$P adopted by Güneş (2014, 2015), or a morpho-syntactic feature, such as the [+comma] feature of Potts (2005).

rayama and Hwang 2019). Studies on object clauses (Ishihara 2020; Hirayama and Hwang 2016) show that clause boundaries do not obligatorily map onto ι-boundaries. It was also argued that Selkirk's (2009) proposal based on FJ *wh*-question data is untenable. Even for the languages that have been reported to exhibit the mapping of embedded clauses to ι (such as German), further examination seems necessary in order to draw a definitive conclusion that the clause-level syntax–prosody mapping exists, because the observed correspondence may have been derived by other factors such as constituent length.

Second, the relevance of speech act to ι-mapping was discussed. It was claimed, following recent proposals (Güneş 2014, 2015; Truckenbrodt 2015), that speech acts, rather than clauses, map to ι. It is proposed that MATCH(illocutionary clause, ι) should be replaced by the mapping of speech acts onto ι's. In place of the clause-level syntax–prosody mapping, an alternative model was suggested in which ι is related to pragmatics–prosody mapping rather than syntax–prosody mapping.

5.2 Clauses as a distinct syntactic category?

Even though a new model of the interfaces between prosody and other modules of grammar was suggested, many questions, both empirical and theoretical, still remain to be further investigated. For example, empirical investigation of the prosody of embedded clauses needs to be continued further, possibly with more elaborated control of other prosodic factors. If speech acts turn out to be relevant for the ι-mapping, integration of speech act theories might be a key to a better understanding of the pragmatics–prosody interface.

Before closing this paper, I would like to mention one fundamental, but still unanswered, theoretical question for future research on syntax–prosody mapping at the clause-level, namely, whether syntactic clauses should be treated as a category distinct from phrases. In standard (generative) syntactic theories, clauses are treated as just a subtype of phrases, either as Complementizer Phrases (CPs) or some other phrases headed by a functional head (e.g., Tense Phrases, TPs). In that sense, clauses are not treated as a syntactic category inherently distinct from other phrases. If clauses are assumed to be a subtype of phrases in syntactic theories, a question immediately arises as to whether a theory of syntax–prosody mapping should treat clauses categorically distinct from phrases, as in Match Theory and Prosodic Adjunction Theory.

In relation to this, it may also be worth noting, as pointed out by Fabian Schubö (p.c.), that phrases that correspond to clauses (ForcePs, CPs, TPs, etc.) are usually considered functional projections rather than lexical projections such as VPs and NPs. This means that, unlike MATCHWORD and MATCHPHRASE, MATCHCLAUSE fails to comply with the Lexical Category Condition (LCC, Selkirk 1996;

Truckenbrodt 1999), which states that "[c]onstraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections" (Truckenbrodt 1999:224). That is, by postulating MatchClause, we would have to assume that the phonological component should be capable of identifying one particular type of functional projection (while ignoring all others) and mark its constituency in the prosodic realization. Such an analysis raises further theoretical questions, such as whether there is any other functional projection that is visible to the phonological component, and if not, why only CPs are visible, etc.

There seem several approaches to this question. For example, Elfner (2012) argues against the LCC. She argues that Match constraints are not sensitive to the distinction between lexical and functional projection but to whether a syntactic projection enters into the phonological component with a phonologically overt material. Another possible approach, as pointed out by an anonymous reviewer, is to reconsider the definition of clausehood, e.g., as proposed by Hamlaoui and Szendrői (2015, 2017), who define a syntactic clause relevant for the syntax–prosody mapping as the "highest projection whose head is overtly filled by the root verb or verbal material" (see also footnote 2). With this definition, clauses can be considered lexical in nature.

Another theoretical question regarding the status of clauses as a distinct category in the syntax–prosody mapping concerns *Layeredness* (Selkirk 1996), or, in Vogel's (2020) term, *Constituent Sequencing*, which is one of the basic principles that constitute what has been called the Strict Layer Hypothesis (SLH, Selkirk 1984, 1986; Nespor and Vogel 1986). The SLH has been considered to characterize the core properties of prosodic structures. While some aspects of the SLH have been weakened in the development of the theory of prosody, such as Exhaustivity (no level-skipping) and Nonrecursivity (no level-repetition), other aspects have still been considered to hold for prosodic hierarchy universally (see, among others, Selkirk 2011; Itô and Mester 2012, 2013; Vogel 2020 for discussion). Layeredness belongs to the latter group. It states that "[a] prosodic constituent of level C_n may only dominate constituents of level C_{n-1} or lower." (Vogel 2020:37). In a prosodic structure that conforms to Layeredness, no ι would be dominated by a φ, for example.

When it comes to syntactic structures, by contrast, it has often been assumed (and often tacitly) that Layeredness is not part of their intrinsic properties, and that the lack of Layeredness in syntax allows the recursion of syntactic structures (as seen in the transformational rules $S \rightarrow NP + VP$; $VP \rightarrow V + S$; $S \rightarrow NP + VP$; ...).

Despite the general assumption that Layeredness is a property unique to prosodic structures, it can, in fact, be said that Layeredness holds in syntax as well, to the extent that morphosyntactic categories, such as morphemes, words, and phrases, seem to be organized in such a way that no category is dominated by a lower category (i.e., morphemes ¡ words ¡ phrases).

There are exceptions to this generalization, namely, the so-called phrasal compounds. (33) shows German examples from Wiese (1996).

(33) Phrasal compounds in German

a. CP: die Wer-war-das-Frage 'the who-was-it question'

b. NP: die Muskel-fur-Muskel-Methode 'the muscle-for-muscle method'

c. AP: der 'Fit-statt-fett'-Buirowettbewerb 'the fit-over-fat office contest'

d. PP: der Zwischen-den-Zeilen-Widerstand 'the between-the-lines opposition'

e. NP: die Gift-in-der-Limonade-Szene 'the poison-in-the-lemonade scene'

(Wiese 1996:184)

As shown in the examples, phrases (and clauses) can be contained in a phrasal compound. However, Wiese (1996) points out that the non-head element of a phrasal compound may be not only a linguistic expression of various syntactic categories, but also an expression from other languages as in (34a), or even a nonverbal expression as in (34b), or just a symbol as in (34c).

(34) a. die No-future-Jugendlichen 'the no-future youngsters'
eine Make-love-not-war-Bewegung 'a make-love-not-war movement'
die Just-in-time-Garantie 'the just-in-time guarantee'

b. seine [nonverbal gesture]-Haltung 'his [nonverbal gesture] attitude'

c. das @-Zeichen 'the @-sign'
die #-Taste 'the #-key'

From these facts, Wiese proposes that in the case of phrasal compounds the non-head element should be treated as a quotation, in which the syntactic category plays no role. In this sense, phrasal compounds are only apparent counterexamples to Layeredness, just as Wiese (1996) argues that they are only apparent counterevidence to the (strong) lexicalist hypothesis, which assumes that the lexicon comprises rules that are necessary and sufficient to define wellformed words. The case of phrasal compounds is somewhat similar to the case of embedded root clauses and parentheticals discussed in §4, in the sense that both non-head elements of phrasal compounds on one side and embedded root clauses and parentheticals on the other exhibit some degree of independence from the superordinate constituent that they are a part of.

Putting the question of phrasal compounds aside, there is another point to be noted regarding Layeredness in syntax. The generalization that Layeredness holds in both prosody and syntax does not

hold between phrases and clauses, as a clause can be embedded inside a phrase (e.g., a VP can contain a clausal object, as in *John [$_{\text{VP}}$ thinks [$_{\text{CP}}$ that Mary is smart]]*.). This exception, however, would be an exception only under the assumption that clauses constitute a distinct category in the syntactic hierarchy above phrases. By contrast, if we assume that clauses are a subtype of phrases, as in standard syntactic theories, the fact that a clause may be dominated by a phrase would not be considered a violation of Layeredness. Put differently, if we give up the idea of the clause-level syntax–prosody mapping (clause $\rightarrow \iota$) and assume that clauses are a subtype of phrases, Layeredness can be maintained as a property of both syntax and prosody alike.

While the theoretical issues regarding the Lexical Category Condition and Layeredness discussed here still await further empirical investigation, it seems evident that there is a fundamental question that still needs to be answered, namely, whether or not it is appropriate to treat clauses as a distinct syntactic category in the theory of the syntax–prosody interface.

Acknowledgments

This study was funded by the Swedish Research Council (2018-01539). I would like to thank the following people for their helpful comments on earlier versions of this paper and/or insightful discussion: Caroline Féry, Chris Golston, Güliz Güneş, Nick Kalivoda, Tomoyuki Kubo, Sara Myrberg, and Fabian Schubö. Special thanks go to the two reviewers, the editors, Junko Ito, Haruo Kubobzono, and Armin Mester, as well as Maho Morimoto, for their comments, which helped improve the content of the paper. Parts of the paper were presented at a few conferences and seminars: the 6th International Conference on Phonetics and Phonology (ICPP 2019) at NINJAL, the 28th Japanese/Korean Linguistics (JK28) at the University of Central Lancashire (online), and ICU Linguistics Colloquium at International Christian University (online). I would like to thank the participants for their questions, comments and discussions. All remaining errors are my own.

References

Beckman, Mary, and Janet Pierrehumbert, (1986). 'Intonational structure in Japanese and English'. *Phonology Yearbook* 3:255–309.

Bianchi, Valentina, and Mara Frascarelli, (2010). 'Is topic a root phenomenon?' *Iberia* 2:43–88.

Chen, Matthew Y., (1987). 'The syntax of Xiamen tone sandhi'. *Phonology Yearbook* 2:109–149.

Deguchi, Masanori, and Yoshihisa Kitagawa, (2002). 'Prosody and wh-questions'. In *NELS 32: Pro-*

ceedings of the North Eastern Linguistics Society, ed. Masako Hirotani, volume 1, 73–92. Amherst, Mass.: GLSA, University of Massachusetts, Amherst.

Dehé, Nicole, (2009). 'Clausal parentheticals, intonational phrasing, and prosodic theory'. *Journal of Linguistics* 45:569–615.

Dehé, Nicole, (2014). *Parentheticals in spoken English*. Cambridge, UK: Cambridge University Press.

Downing, Bruce, (1970). 'Syntactic structures and phonological phrasing in English'. Doctoral Dissertation, University of Texas, Austin.

Ebert, Cornelia, (2009). *Quantificational topics: A scopal treatment of exceptional wide scope phenomena*. Dordrecht: Springer.

Elfner, Emily Jane, (2012). 'Syntax–prosody interaction in Irish'. Doctoral Dissertation, University of Massachusetts, Amherst.

Feldhausen, Ingo, (2010). *Sentential form and prosodic structure of Catalan*. Amsterdam/Philadelphia: John Benjamins.

Frascarelli, Mara, (2000). *The syntax-phonology interface in focus and topic constructions in Italian*. Dordrecht: Kluwer.

Fujisaki, Hiroya, and Keikichi Hirose, (1984). 'Analysis of voice fundamental frequency contours for declarative sentences of Japanese'. *Journal of the Acoustical Society of Japan* 5:233–242.

Fujisaki, Hiroya, and Hiroshi Sudo, (1971). 'A model for the generation of fundamental frequency contours of Japanese word accent'. *Journal of Acoustic Society of Japan* 27:445–453.

Grijzenhout, Janet, and Baris Kabak, ed., (2009). *Phonological domains: Universals and deviations*. Berlin/New York: Mouton de Gruyter.

Güneş, Güliz, (2014). 'Constraints on syntax-prosody correspondence: The case of clausal and subclausal parentheticals in Turkish'. *Lingua* 150:278–314.

Güneş, Güliz, (2015). 'Deriving prosodic structures'. Doctoral Dissertation, University of Groningen.

Hamlaoui, Fatima, and Kriszta Szendrői, (2015). 'A flexible approach to the mapping of intonational phrases'. *Phonology* 32:79–110.

Hamlaoui, Fatima, and Kriszta Szendrői, (2017). 'The syntax-phonology mapping of intonational phrases in complex sentences: A flexible approach'. *Glossa: a journal of general linguistics* 1–31.

Hayata, Teruhiro, (1985). *Hakata hōgen no akusento / keitairon [the accent and morphology of the Hakata dialect]*. Fukuoka: Kyushu University Press.

Hirayama, Manami, and Hyun Kyung Hwang, (2016). 'Not all XPs affect prosody in Japanese'. In *NELS 46: Proceedings of the North Eastern Linguistics Society*, ed. Christopher Hammerly and Brandon Prickett, 95–104.

Hirayama, Manami, and Hyun Kyung Hwang, (2019). 'Relative clause and downstep in Japanese'. In *Supplemental Proceedings of the 2018 Annual Meeting on Phonology*, ed. Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose, and Matt Zaslansky. Washington, DC: Linguistic Society of America.

Hirose, Yuki, and Yoshihisa Kitagawa, (2011). 'Production–perception asymmetry in wh-scope marking'. In *Processing and producing head-final structures*, ed. Hiroko Yamashita, Yuki Hirose, and Jerome L. Packard, chapter 5, 93–110. Dordrecht: Springer.

Hirotani, Masako, (2005). 'Prosody and LF: Processing Japanese wh-questions'. Doctoral Dissertation, University of Massachusetts, Amherst.

Hwang, Hyun Kyung, (2011). 'Scope, prosody, and pitch accent: The prosodic marking of wh-scope in two varieties of Japanese and South Kyeongsang Korean'. Doctoral Dissertation, Cornell University.

Igarashi, Yosuke, (2015). 'Intonation'. In Kubozono (2015), chapter 13, 525–568.

Igarashi, Yosuke, (2018). 'Intonation'. In *The Cambridge handbook of Japanese linguistics*, ed. Yoko Hasegawa, chapter 9, 181–201. Cambridge, UK: Cambridge University Press.

Inkelas, Sharon, and Draga Zec, (1995). 'Syntax-phonology interface'. In *The handbook of phonological theory*, ed. John A. Goldsmith, 535–549. Oxford, UK: Blackwell.

Ishihara, Shinichiro, (2002). 'Invisible but audible wh-scope marking: Wh-constructions and deaccenting in Japanese'. In *Proceedings of the 21st West Coast Conference on Formal Linguistics (WCCFL 21)*, ed. Line Mikkelsen and Chris Potts, 180–193. Somerville, Mass.: Cascadilla Press.

Ishihara, Shinichiro, (2003). 'Intonation and interface conditions'. Doctoral Dissertation, Massachusetts Institute of Technology. URL `http://hdl.handle.net/1721.1/17020`.

Ishihara, Shinichiro, (2015). 'Syntax–phonology interface'. In Kubozono (2015), chapter 14, 569–618.

Ishihara, Shinichiro, (2016). 'Japanese downstep revisited'. *Natural Language & Linguistic Theory* 34:1389–1443. URL `http://dx.doi.org/10.1007/s11049-015-9322-8`.

Ishihara, Shinichiro, (2020). 'Testing MatchClause: The prosody of embedded clauses in Japanese'. Paper presented at the 28th Japanese/Korean Linguistics.

Itô, Junko, and Armin Mester, (2007). 'Prosodic adjunction in Japanese compounds'. In *Formal Approaches to Japanese Linguistics 4 (FAJL4)*, 97–111.

Itô, Junko, and Armin Mester, (2012). 'Recursive prosodic phrasing in Japanese'. In *Prosody matters: Essays in honor of Elisabeth Selkirk*, ed. Toni Browsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara, 280–303. London: Equinox Publishing.

Itô, Junko, and Armin Mester, (2013). 'Prosodic subcategories in Japanese'. *Lingua* 124:20–40.

Jacobs, Joachim, (1984). 'Funktionale Satzperspektive und Illokutionssemantik'. *Linguistische Berichte* 91:25–58.

Kawahara, Shigeto, and Takahito Shinya, (2008). 'The intonation of gapping and coordination in Japanese: Evidence for intonational phrase and utterance'. *Phonetica* 65:62–105.

Kentner, Gerrit, and Caroline Féry, (2013). 'A new approach to prosodic grouping'. *The Linguistic Review* 30:277–311.

Kishimoto, Hideki, (2001). 'Binding of indeterminate pronouns and clause structure in Japanese'. *Linguistic Inquiry* 32:597–663.

Kitagawa, Yoshihisa, and Janet Dean Fodor, (2003). 'Default prosody explains neglected syntactic analysis in Japanese'. In *Japanese/Korean Linguistics*, ed. William McClure, volume 12, 267–279. Stanford: CSLI Publications.

Kitagawa, Yoshihisa, and Janet Dean Fodor, (2006). 'Prosodic influence on syntactic judgements'. In *Gradience in grammar: Generative perspective*, ed. Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky, and Ralf Vogel, chapter 17. Oxford University Press.

Kitagawa, Yoshihisa, and Yuki Hirose, (2012). 'Appeals to prosody in wh-interrogatives: Speakers' versus listeners' strategies'. *Lingua* 122:608–641.

Krifka, Manfred, (2001). 'Quantifying into question act'. *Natural Language Semantics* 9:1–40.

Kubo, Tomoyuki, (1989). 'Fukuoka-shi hōgen no dare, nani tō no gimon-shi o fukumu bun no pitchi patān [The pitch pattern of the sentences containing interrogative markers such as dare, nani, etc. in Fukuoka dialect]'. *Kokugo-gaku* 156:71–82.

Kubo, Tomoyuki, (2001). 'Fukuoka hōgen ni okeru tōgoron to on'inron no kyōkai ryōiki. [Syntax–phonology interface in the Fukuoka dialect.]'. *Journal of the Phonetic Society of Japan* 5:27–32.

Kubo, Tomoyuki, (2005). 'Phonology–syntax interfaces in Busan Korean and Fukuoka Japanese'. In *Proceedings from the Symposium Cross-linguistic Studies of Tonal Phenomena: Historical Development, Tone-Syntax Interface, and Descriptive Studies*, ed. Shigeki Kaji, 195–219. Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

Kubozono, Haruo, (1993). *The organization of Japanese prosody*. Tokyo: Kurosio Publishers.

Kubozono, Haruo, ed., (2015). *Handbook of Japanese phonetics and phonology*, volume 2 of *Handbooks of Japanese Language and Linguistics*. Berlin / Boston: Mouton de Gruyter.

Kuroda, S.-Y., (1965). 'Generative grammatical studies in the Japanese language'. Doctoral Dissertation, Massachusetts Institute of Technology.

Kuroda, S.-Y., (2013). 'Prosody and the syntax of indeterminates'. *Lingua* 124:64–95.

Ladd, D. Robert, (1986). 'Intonational phrasing: the case for recursive prosodic structure'. *Phonology Yearbook* 3:311–340.

Ladd, D. Robert, (1988). 'Declination 'reset' and the hierarchical organization of utterance'. *Journal of the Acoustic Society of America* 84:530–544.

Maekawa, Kikuo, (1991). 'Tōkyō hōgen gimon-bun no intonēsyon [The intonation of question sentences in Tokyo dialect]'. In *Proceedings of the 1991 PSJ Annual Convention*, 42–47. The Phonetic Society of Japan.

Maekawa, Kikuo, (1997). 'Nihongo gimonsi gimon-bun no intonēsyon [The intonation of the Japanese wh-question]'. In *Bunpō to onsei [Grammar and sound]*, ed. Onsei Bunpō Kenkyū-kai, chapter 3, 45–53. Tokyo: Kurosio Publishers.

Maekawa, Kikuo, Hideaki Kikuchi, Yosuke Igarashi, and Jennifer Venditti, (2002). 'X-JToBI: An extended J_ToBI for spontaneous speech'. In *Proceedings of The 7th International Conference on Spoken Language Processing (ICSLP)*, 1545–1548. Denver, Colorado.

McCawley, James D., (1965). 'The accentual system of standard Japanese'. Doctoral Dissertation, Massachusetts Institute of Technology.

Myrberg, Sara, (2010). 'The intonational phonology of Stockholm Swedish'. Doctoral Dissertation, Stockholm University.

Myrberg, Sara, (2013). 'EqualSysters and prosodic branching'. *Phonology* 30:73–124.

Nelson, Gerald, Sean Wallis, and Bas Aarts, (2002). *Exploring natural language: Working with the british component of the international corpus of english.*. Amsterdam: John Benjamins.

Nespor, Marina, and Irene Vogel, (1986). *Prosodic phonology*. Dordrecht: Foris Publications.

Nishigauchi, Taisuke, (1990). *Quantification in the theory of grammar*. Dordrecht: Kluwer Academic Publishers.

Pierrehumbert, Janet, and Mary Beckman, (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.

Poser, William J., (1984). 'The phonetics and phonology of tone and intonation in Japanese'. Doctoral Dissertation, Massachusetts Institute of Technology.

Potts, Christopher, (2005). *The logic of conversational implicature*. Oxford University Press.

Rizzi, Luigi, (1997). 'The fine structure of the left periphery'. In *Elements of grammar: Handbook in generative syntax*, ed. Liliana Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Schubö, Fabian, (2020). 'Intonation Phrase formation in sentences with clausal embedding'. *Journal of Comparative Germanic Linguistics* 23:161–201.

Selkirk, Elisabeth, (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth, (1986). 'On derived domains in sentence prosody'. *Phonology Yearbook* 3:371–405.

Selkirk, Elisabeth, (1996). 'The prosodic structure of function words'. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James L. Morgan and Katherine Demuth, 187–213. Mahwah, N.J.: Lawrence Erlbaum Associates.

Selkirk, Elisabeth, (2005). 'Comments on intonational phrasing in English'. In *Prosodies: With special reference to Iberian languages*, ed. Sónia Frota, Mariana Vigário, and Maria Jo ao Freitas, 11–59. Berlin/New York: Mouton de Gruyter.

Selkirk, Elisabeth, (2009). 'On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure'. *Gengo Kenkyu* 136:35–74.

Selkirk, Elisabeth, (2011). 'The syntax-phonology interface'. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, 435–484. Oxford, UK: Blackwell, 2nd edition.

Selkirk, Elisabeth, and Koichi Tateishi, (1988). 'Constraints on minor phrase formation in Japanese'. In *CLS*, ed. Lynn MacLeod, Gary Larson, and Diane Brentari, volume 24 Part One: The General Session, 316–336. Chicago, IL: Chicago Linguistic Society.

Shimoyama, Junko, (2001). 'Wh-constructions in Japanese'. Doctoral Dissertation, University of Massachusetts, Amherst.

Shinya, Takahito, Elisabeth Selkirk, and Shigeto Kawahara, (2004). 'Rhythmic boost and recursive minor phrase in Japanese'. In *Proceedings of the Second International Conference on Speech Prosody*, 345–348.

Smith, Jennifer L., (2005). 'On the WH-question intonational domain in Fukuoka Japanese: Some implications for the syntax-prosody interface'. In *Papers on prosody*, ed. Shigeto Kawahara, volume 30 of *UMOP*, 219–237. Amherst, Mass.: GLSA.

Smith, Jennifer L., (2013). 'Fukuoka Japanese wh prosody in production and perception'. *Lingua* 124:96–130.

Smith, Jennifer L., (2016). '*Wh* prosody is not focus prosody in Fukuoka Japanese'. In *Japanese/Korean Linguistics 23*, ed. Michael Kenstowicz, Ted Levin, and Ryo Masuda, 23–37. Standord: CSLI Publications.

Truckenbrodt, Hubert, (1999). 'On the relation between syntactic phrases and phonological phrases'. *Linguistic Inquiry* 30:219–255.

Truckenbrodt, Hubert, (2005). 'A short report on intonation phrase boundaries in German'. *Linguistische Berichte* 203:273–296.

Truckenbrodt, Hubert, (2015). 'Intonation phrases and speech acts'. In *Parenthesis and ellipsis: Cross-linguistic and theoretical perspective*, ed. Marlies Kluck, Dennis Ott, and Mark de Vries, 301–349. Berlin/Boston: De Gruyter Mouton.

Truckenbrodt, Hubert, and Isabelle Darcy, (2010). 'Object clauses, movement, and phrasal stress'. In *The sound patterns of syntax*, ed. Nomi Erteshik-Shir and Lisa Rochman, 189–216. Oxford/New York: Oxford University Press.

Truckenbrodt, Hubert, and Caroline Féry, (2015). 'Hierarchical organisation and tonal scaling'. *Phonology* 32:19–47.

Uyeno, Tazuko, Hideo Hayashibe, and Kunihiko Imai, (1979). 'On pitch contours of declarative, complex sentences in Japanese'. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics* 13:175–187.

Uyeno, Tazuko, Hideo Hayashibe, Kunihiko Imai, Hiroshi Imagawa, and Shigeru Kiritani, (1981). 'Syntactic structure and prosody in Japanese: A study on pitch contours and the pauses at phrase boundaries'. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics* 15:91–108.

Venditti, Jennifer J., (2005). 'The J_ToBI model of Japanese intonation'. In *Prosodic typology: The phonology and intonation of phrasing*, ed. Sun-Ah Jun, 172–200. Oxford, UK: Oxford University Press.

Venditti, Jennifer J., Kikuo Maekawa, and Mary E. Beckman, (2008). 'Prominence marking in the Japanese intonation system'. In *The oxford handbook of Japanese linguistics*, ed. Shigeru Miyagawa and Mamoru Saito, 456–512. Oxford/New York: Oxford University Press.

Vogel, Irene, (2009). 'The status of the Clitic Group'. In *Phonological domains: Universals and deviations*, ed. Janet Grijzenhout and Baris Kabak, 15–46. Berlin/New York: Mouton de Gruyter.

Vogel, Irene, (2020). 'Life after the Strict Layer Hypothesis: Prosodic structure geometry'. In *Prosodic studies: Challenges and prospects*, ed. Hongming Zhang and Youyong Qian, chapter 1, 10–60. London: Routledge.

de Vries, Mark, (2012). 'Unconventional mergers'. In *Ways of structure building*, ed. Myriam Uribe-Etxebarria and Vidal Valmala, 143–166. Oxford, UK: Oxford University Press.

Wagner, Michael, (2005). 'Prosody and recursion'. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Mass.

Wiese, Richard, (1996). 'Phrasal compounds and the theory of word syntax'. *Linguistic Inquiry* 27:183–193. URL `https://www.jstor.org/stable/4178931`.

Zec, Draga, and Sharon Inkelas, (1990). 'Prosodically constrained syntax'. In *The phonology-syntax connection*, ed. Sharon Inkelas and Draga Zec, 365–378. Stanford, California: CSLI Publications.