Original Article

second language research

Second Language Research
1–31

DOI: 10.1177/0267658319897786
journals.sagepub.com/home/slr
SAGE

Acquisition of L2 Japanese WH questions: Evidence of phonological contiguity and non-shallow structures

John Archibald and Nicole Croteau
University of Victoria, Canada

Abstract

In this article we look at some of the structural properties of second language (L2) Japanese WH questions. In Japanese the WH words are licensed to remain *in situ* by the prosodic contiguity properties of the phrases which have no prosodic boundaries between the WH word and the question particle. In a rehearsed-reading, sentence production task, we look to see whether non-native speakers of Japanese who are learning the L2 in university classes in North America are able to acquire grammars which are constrained by such universal properties as Match Theory and Contiguity Theory. While linear mixed effects analyses of the pitch contours reveal that the L2ers have not acquired the phonetic implementation distinction of the documented pitch boost on WH words compared to non-WH DPs, our data show that the participants have acquired the pitch compression patterns indicative of having no prosodic phrases intervening between the WH word and the question particle. This property of Japanese WH questions is not taught in their classes, and, thus we argue, that the data are supportive of the position that interlanguage grammars are constrained by universal grammatical properties such as the prosodic contiguity of WH-phrase licensing. We also present these results as being counter to the Shallow Structure Hypothesis.

Keywords

Japanese, phonetics/phonology interface, phonology, phonology–syntax interface, shallow structure

I Introduction

The field of second language acquisition (SLA) looks to probe the properties of the grammars of second language (L2) learners. While there is much research that investigates a single module such as phonology (Archibald, 1998; Cabrelli Amaro, 2017) or syntax (White, 2003), there has also been much written on the grammatical interfaces (Goad and White, 2004; Montrul, 2011; Sorace, 2011; White, 2011). Considerable attention has been given to the morphology/syntax interface (Franceschina, 2001), and also to the syntax–pragmatics

Corresponding author:
John Archibald, University of Victoria, Clearihue D271, Victoria, BC, V8W 2Y2, Canada.
Email: johnarch@uvic.ca

interface (Sorace and Filiaci, 2006), but there has been less attention given to the phonology–syntax interface (see Fodor, 2002) in SLA, though, of course, it has been addressed in the theoretical literature (Elfner, 2015). This article looks at one aspect of the phonology–syntax interface related to the formation of WH questions.

Richards (2010, 2016) articulates a theory which accounts for the variation between (a) languages that move WH elements to the left periphery, and (b) languages that allow WH elements to stay in place. English generates WH questions via WH movement as shown in (1). Japanese, on the other hand, allows the WH word to remain *in situ*, as shown in (2).

(1) *Whom* should Bob call?

(2) Mito-ga *nani*-o katta no?
Mito-NOM what-ACC bought +Q
'What did Mito buy?'

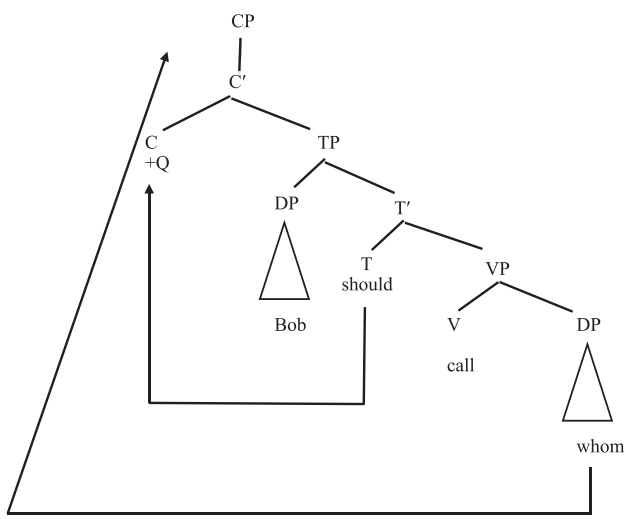
In this article, we investigate the second language acquisition of Japanese WH questions in order to explore whether the interlanguage grammars of the learners are subject to the grammatical principles laid out in the syntactic and phonological literature.

II The syntax–phonology interface

1 Contiguity theory

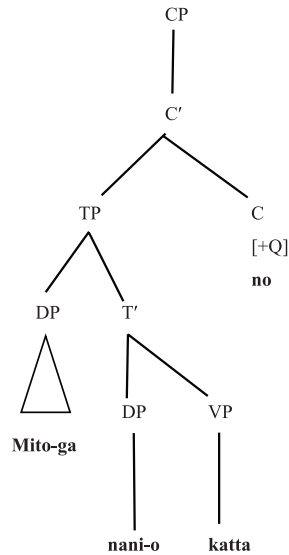
Richards (2010, 2016) argues that these are two strategies to achieve the contiguity of the WH word and the interrogative feature [+Q]. In English, surface linear adjacency achieves contiguity as a result of moving the WH word to the Spec of CP position at the left edge (adjacent to where the phonologically null [+Q] feature resides), as shown in (3).

(3)



In Japanese, there are two noteworthy differences from the English structure, as we see in (4), 'What did Mito buy?'

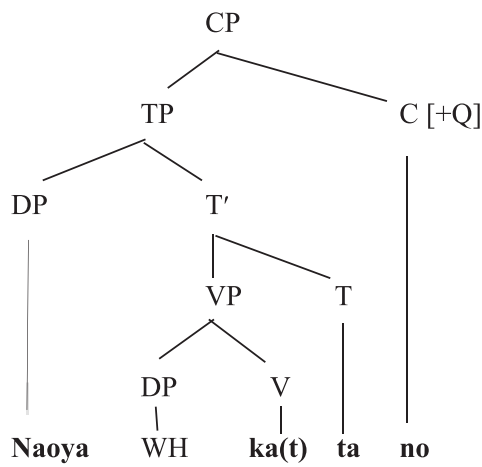
(4)



First, we note that the [+Q] feature is spelled out by a question particle (which in (4) is *no*[1]). Second, the WH word does not move, to the periphery (either left or right) but contiguity is achieved in another fashion: prosodic contiguity. Richards demonstrates that (a) there are no prosodic boundaries between the WH word (in this case *nanio*) and the [+Q] feature, and (b) there is what he calls a phonetic boost on the WH phrases (compared to non-WH DPs).

Richards draws on work on the prosody of Japanese WH questions (Deguchi and Kitagawa, 2002; Hirotani, 2005; Ishihara, 2003; Sughara, 2003). He provides pitch tracks of the Tokyo dialect which show a pitch boost on the WH element (like *nanio*) compared to DP direct objects (like *nanika*). Our pitch tracks are almost identical to those given in Richards. The structure created by Merge[2] would be (Japanese being right-headed) as shown in (5).

(5) 'What did Naoya buy?'



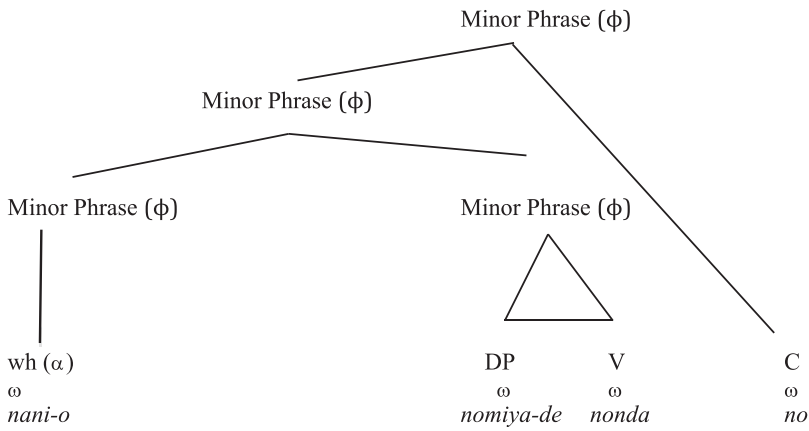
More schematically, the constituent structure would be that in (6). The strategy to achieve prosodic contiguity that Japanese invokes is to eliminate the prosodic boundaries between the WH phrase and the Complementizer, and generate the prosodic structure given in (7).

(6) [DP] [[$_{DP}$ WH] [V] $_{VP}$] C

(7) (φ DP) (φ WH C)

Let us look at how this plays out when there are more elements between the WH phrase and the Complementizer. Richards (2010) looks at Minor Phrases (which dominate prosodic words). He follows Kubozono (2006) in the analysis that Japanese Minor Phrases can be recursive; multiple Minor Phrases are dominated by an overarching Minor Phrase, as shown in (8).

(8) 'drank what at the bar?'



He refers to the highest Minor Phrase as the WH domain. Given a WH phrase α and a complementizer C where α takes scope, α and C must be separated by as few Minor Phrase boundaries as possible, for a given level of Minor Phrasing.[3]

This phonological phrasing would be generated in a fashion consistent with Match Theory (Elfner, 2015; Selkirk, 2011). Match Theory probes the syntax–phonology interface. In attempting to describe the domains of certain phonological phenomena, sometimes syntactic structures appear to be relevant. For example, English phrasal stress falls on the rightmost element in a syntactic phrase (Chomsky and Halle, 1968) as shown in (9).

(9) [A syntactícian] and [a phonólogist] [walk into a bár]

Yet, other phonological phenomena are not predictable by syntactic structures, but rather by prosodic domains. Some patterns apply at the highest level of phonological utterance, such as RP-English r-insertion (Nespor and Vogel, 1986) which applies when there are adjacent vowels which can be adjacent even across sentences, as shown in (10).

(10) neve[r] again
the giant Panda[r], as you know, is an endangered species
That's a nice ca[r]. Is it yours?

Other changes, such as Tuscan spirantization (where stops become fricatives) are somewhat more restricted insofar as parenthetic material blocks the change. The domain of application, then, is the intonational phrase, as shown in (11). The orthographic 'c' words which are underscored are produced as the fricative [h] while the orthographic 'c' words which are not underscored are produced as the stop [k].

(11) i canarini congolesi costano molto cari
'Congolese canaries are very expensive.'

i canarini congolesi, come sai, costano molto cari
'Congolese canaries, as you know, are very expensive.'

English stress retraction occurs within a phonological phrase but not across the boundary of two phonological phrases. The sentence in (12a) shows stress retraction while (12b) does not.

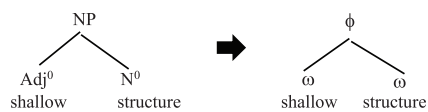
(12) a. [pérseveres gladly] (compare with persevéres)
b. persevéres] [gládly and diligently]

Recent research has demonstrated that there is a systematic correspondence between syntactic and prosodic structure. These correspondences are articulated in Match Theory.

2 Match theory

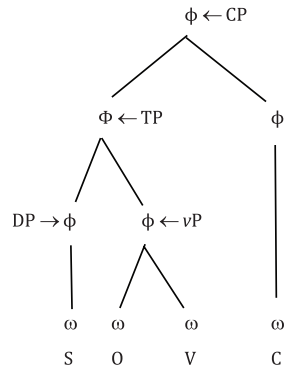
Richards' work is couched within the larger literature of the syntax–prosody interface (e.g. Selkirk, 2011; Tyler, 2019). Match Theory, broadly speaking, begins with the assumption of syntax–prosody isomorphism. Phonological structure is modelled via the Prosodic Hierarchy (Nespor and Vogel, 1986) with modifications to earlier proposals which recognize that (a) not all languages may encode all levels (see Özcelik, 2017) and (b) that recursive structures are permissible. The theory accounts for the correspondences between phonological and syntactic structures. The preferential mapping is between (a) syntactic phrases (XPs) and phonological phrases (ϕ), and (b) syntactic heads (X^0s) and prosodic words (ω).[4] Furthermore, it is assumed that CPs are mapped on to intonational phrases (ι). Within such a model (e.g. Tyler, 2019), an NP such as *shallow structure* would have the structural mapping shown in (13) at spell out.[5]

(13)



This transparent mapping between the syntax and the phonology which would generate the following structures for the sentences we are investigating (where the arrows indicate the mapping between the syntactic structure and the phonological structure given in (14).

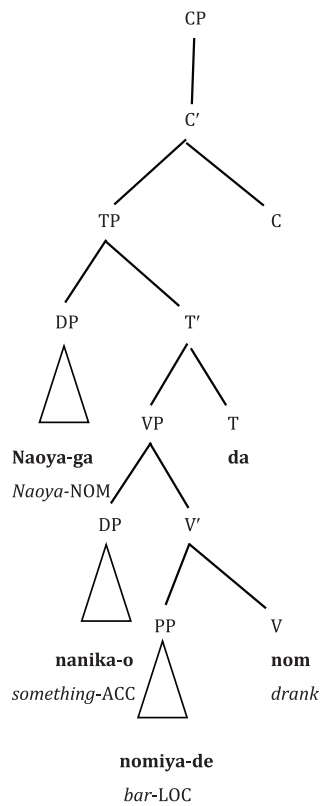
(14)



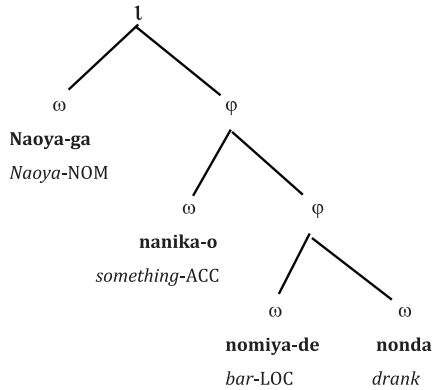
Thus, the prosody preserves the hierarchical syntactic structure. For the structure to be well-formed the *in situ* WH phrase must be followed by a minimal number of intervening φ-phrases.

Let us look at a specific example from Richards (2016). We begin with the declarative sentence, 'Naoya drank something in the bar' shown in (15). This syntactic structure would be mapped on the prosodic tree shown in (16).

(15)

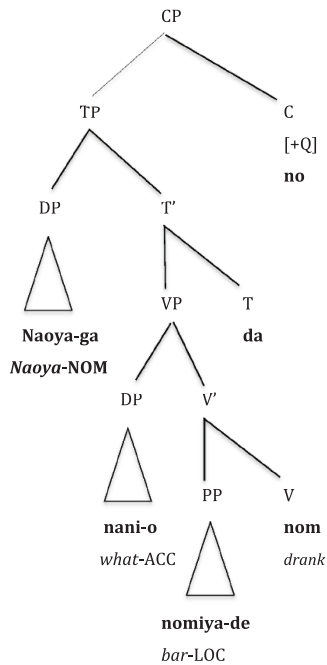


(16)



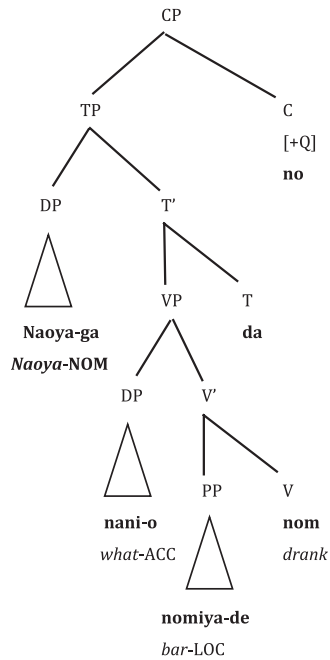
Richards argues that such prosodic trees[6] are generated by the narrow syntax as the derivation proceeds, not, as is standardly assumed after the narrow-syntactic derivation is finished. Thus, abstract phonology is not merely a late spell-out routine. Now let us consider the corresponding interrogative sentence 'What did Naoya drink at the bar?', shown in (17).

(17)



The generation of the prosodic tree proceeds through several intermediate steps which need not concern us here to arrive at the structure given in (18).

(18)



It is worth noting here some of Richards' (2010, 2016) assumptions about the interaction of the syntax with the phonology. Note that he assumes that WH movement takes place only if the prosody requires it. While he maintains that WH movement is a syntactic, not a late phonetic, operation, he is agnostic as to the mechanism which allows the syntax to 'look ahead' to the phonology.

3 What is to be acquired?

The research question we seek to investigate in this article is:

> Are advanced L2 speakers able to set up a phonological grammar with no prosodic boundaries between the WH word and the Question complementizer ([+Q]) in order to properly license WH *in situ* as would be the case if Richards' (2016) Contiguity Theory holds for interlanguage grammars?

In order to understand the operationalization of this question, let us compare two sentences[7] given in (19) and (20), where the direct objects are in bold, and the minor phonological phrases are in italics.

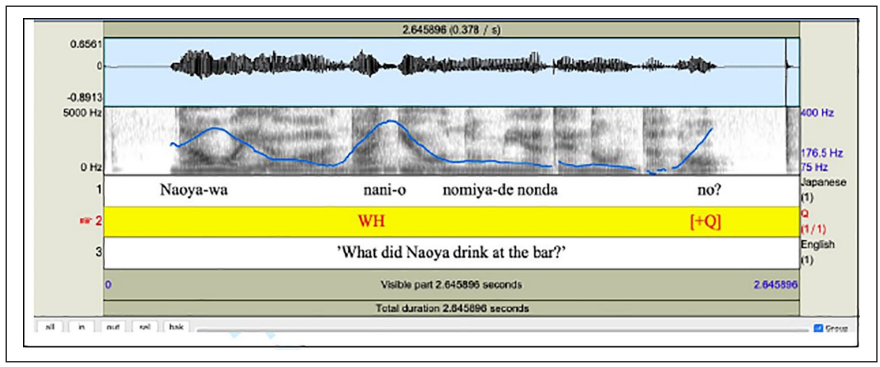


Figure 1. Pitch track of the WH question shown in (19).

(19) Naoya-wa **nani**-o *nomiya-de* *nonda* no?
Naoya-TOP what-ACC bar-LOC drank +Q
ナオヤは、何を飲み屋で飲んだの？
'What did Naoya drink at the bar?'

(20) Naoya-ga **nanika**-o *nomiya-de* *nonda*.
Naoya-NOM something-ACC bar-LOC drank
ナオヤが、何かを飲み屋で飲んだ。
'Naoya drank something at the bar.'

For native speakers of Japanese, we expect higher pitch on the objects which are WH words – such as *nanio* in (19) – compared with objects which are DPs, such as *nanika* in (20). Readers should note that we use the label DP (recognizing that WH words are, in fact, found within DPs) to refer to DPs which lack a WH word thus allowing us to compare the behaviour of WH phrases with DPs. The spectrograms in Figures 1 and 2 are taken from one of the native speaker controls in this study (a female voice is shown). Statistical analyses will be reported in section 3 but for now a visual comparison will suffice. Note the pitch spike on the WH phrase (*nanio*). Compare this with the declarative sentence of (20) produced by the same speaker shown in Figure 2.

Note the lower pitch (209 Hz) on the DP object (*nanika*), which is at the intersection of the red lines. These pitch tracks are virtually identical to those given in Richards (2010: 144). Given this convergence, we assume that our production task is revealing the same properties of Japanese pitch as Ishihara (2003), and we, thus, assume that native speakers of Japanese implement this pitch boost on WH phrases. Furthermore, we find no pitch peaks on the prosodic words (e.g. *nomiya de nonda*) which come between the WH word and the [+Q] particle; as in (19).[8] The pitch tracks in Figures 3 and 4 are taken from one of our native speaker controls.[9] This is in contrast to the pitch peak we see in the declarative sentence in Figure 4.

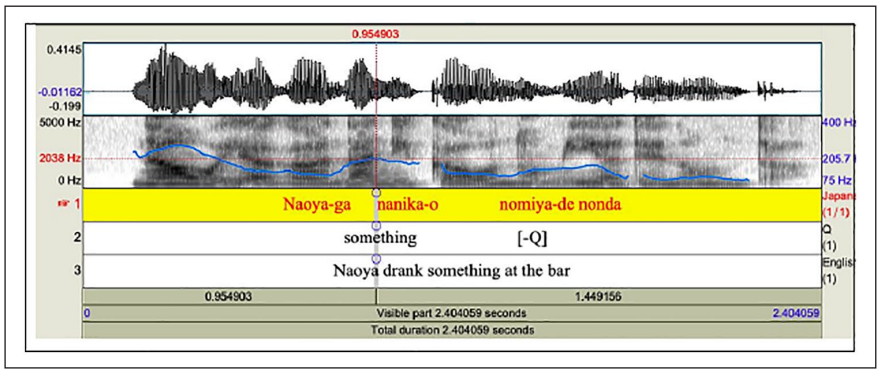


Figure 2. Pitch track of declarative sentence of (20) with DP object.

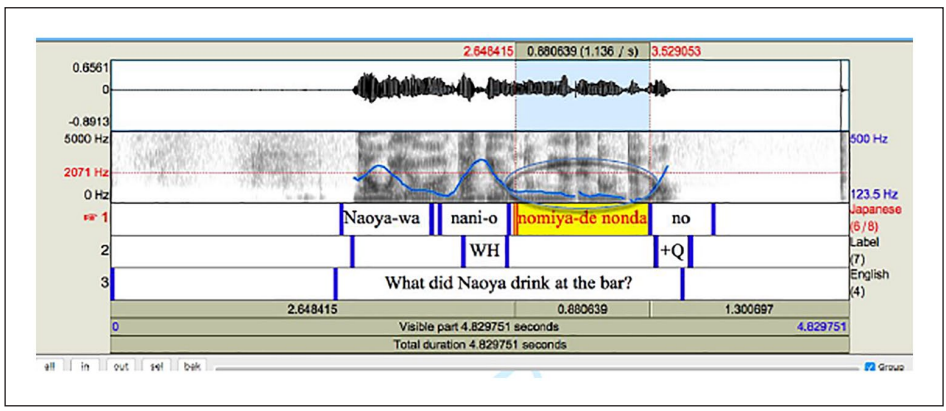


Figure 3. A native speaker recorded in our study demonstrating pitch compression.

III Methodology

In this study, we looked to see whether non-native speakers (NNS) of Japanese are able to acquire L2 grammars which observe this property of phonological contiguity. Sixteen instructed learners of Japanese (9 self-assessed advanced proficiency; 7 self-assessed intermediate proficiency) were audio recorded. In addition, there were 6 native speakers of Japanese to act as controls in a subsequent analysis of prosodic contiguity.

The participants[10] were recruited as volunteers from announcements in senior-level Japanese classes at a Canadian university.[11] The advanced group consisted of 4 males and 5 females while the intermediate group consisted of 4 females and 3 males. Twelve were native speakers of English while 4 were native speakers of other languages (Cantonese, Mandarin, or Korean). Furthermore, this heterogeneity of first language (L1), in our view, is not a major impediment insofar as we are not looking to explore any aspects of L1 transfer but rather whether the L2 production is nativelike with respect to the phonological and phonetic properties of the WH domain. The participant pool is also

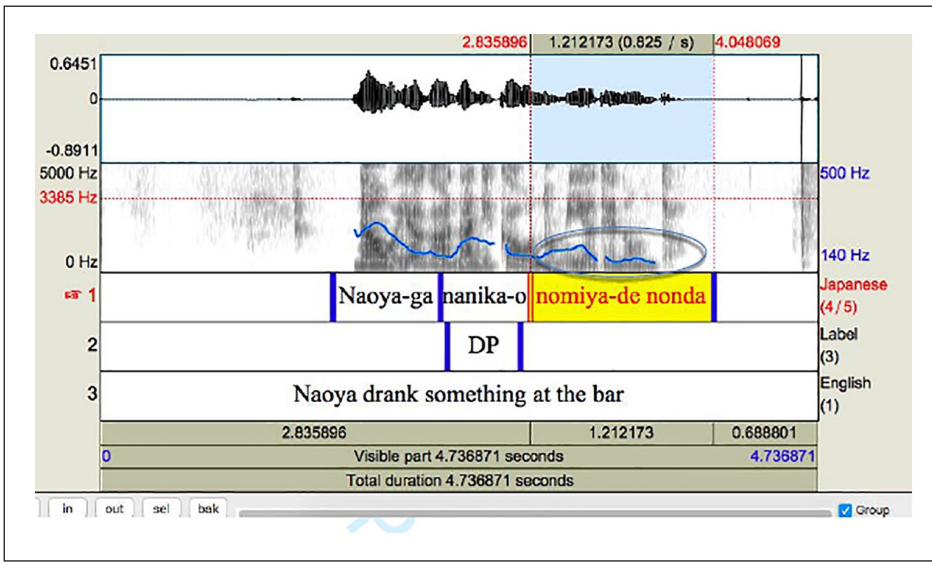


Figure 4. Native speaker pitch rise in declarative sentence.

mixed in that there are both males and females. We have grouped them together (after transforming the data) for analysis because nothing in the analysis depends on the absolute pitch of the production. Rather, it is the pitch change that is of interest. All analyses were performed in R version 3.2.3 (R Core Team, 2015) with the package 'lme4' version 1.1-12 (Bates, 2015).

1 The task

The participants were given in advance 19 Japanese sentences (see Appendix 1[12]) which they were asked to rehearse at home in preparation for reading aloud assuming a neutral discourse context. If, during recording in the lab, they made a slip, they were allowed to record again. The vast majority of sentences were read only once. Sometimes participants requested a second attempt. To us, this is an indication that the task was feasible, and not too demanding on the learners. After they finished a sentence, it was played back for them, and were asked if they wished to record it again. If they did not, we proceeded to the next sentence. The final version was the one used for acoustic analysis. In this way, we tried to ensure that the production we analysed was the best possible production that the participant could produce. Two examples are given in (21) and (22).[13]

(21)	Miki-wa	kinō	nani-o	kai-mashi-ta	ka?
	Miki-TOP	yesterday	what-ACC	buy-POL-PAST	+Q
	ミキは、昨日何を買いましたか？				
	'What did Miki buy yesterday?'				

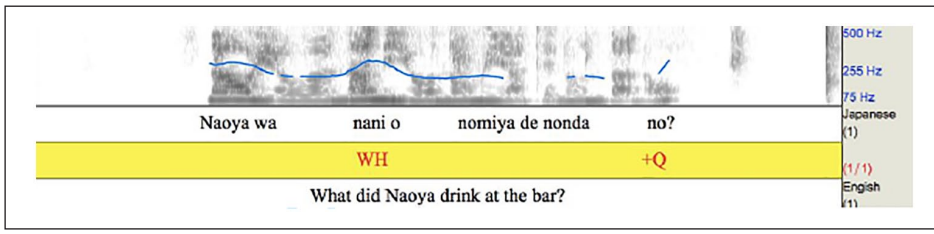


Figure 5. Visualization of the pitch patterns.

(22) Miki-wa kinō hon-o kai-mashi-ta.
Miki-TOP yesterday book-ACC buy-POL-PAST
ミキは、昨日本を買いました。
'Miki bought a book yesterday.'

Sentences were recorded on Audacity at 44.1 kHz, and pitch tracking was done via Praat 6.0.09 (Boersma and Weenink, 2019) after exporting the .wav files from Audacity. An example of the pitch tracking is shown in Figure 5, where we note the pitch rise on *nanio* and *no*.

The pitch peak was measured and entered into a spreadsheet for calculation. Given that we were comparing different lexical items, we controlled for pitch accent in the stimuli. All of the target words (both WH and DP) contained both high and low pitch accents, so any difference which could potentially emerge in the average across participants and across lexical items cannot be attributed to some words having either all high or all low pitch accents.[14]

2 The structure of the argument

Native speakers of Japanese boost the pitch on WH words (compared to non-WH DPs) and compress the pitch between the WH word and the Complementizer. Richards (2016) builds a theory in which (a) the WH *in situ* is licensed by prosodic contiguity, and (b) the phonology–syntax interface is mediated by an architecture following Match Theory (Elfner, 2015; Selkirk, 2011), which predicts that the phonological and syntactic constituents should be largely isomorphic. Our study is designed to investigate these properties in second language learners. If the L2 participants are producing prosodic contours which are nativelike in that they reflect L2 Japanese syntax, this is something that they could not have learned easily from classroom input. We have given our participants a reading task (not a grammaticality or acceptability judgement task), and make the assumption (Levelt, 1989) that the L2ers are generating a syntactic representation in production. So, given the syntactic structure of WH *in situ*, we are looking to see if the produced utterances reflect nativelike prosody. There is much literature which demonstrates that syntactic constituent structure is invoked in silent reading tasks (e.g. Juffs and Harrington, 1998; Kim et al., 2015; Pliatsikas and Marinis, 2013), and we maintain this is true in production tasks as well. There are four main reasons why this is a question of interest:

1. It will add to the literature exploring whether interlanguage grammars are constrained by the same principles as primary languages.
2. It will add to the literature exploring whether non-native speakers can attain L2 grammatical properties which are (a) not taught in class, and (b) not easily read off the input.
3. It will provide further empirical data relevant to the Shallow Structure Hypothesis (Clahsen and Felser, 2017) which argues that second language learners often have problems processing abstract hierarchical grammatical constituents in real time.
4. It will provide data relevant to the question of whether L2 phonetic ability and L2 phonological knowledge proceed along the same developmental path.

Let us examine each of these contributions in a bit more depth.

a Interlanguages as natural languages. This is a thread which has been addressed often in the literature (Bley-Vroman, 2009) under a number of guises but it really comes down to the issue of domain specificity. Many researchers have probed the question of whether interlanguage grammars are constrained by universal linguistic principles (White, 2003). Of particular interest are the questions of whether these universal properties are evident when there is no overt evidence for them in the L1. Will, for example, an L1-speaker who has no overt WH movement show knowledge of constraints on WH movement in the interlanguage? These poverty of the stimulus cases are summarized nicely in Özcelik and Sprouse, 2016; Özcelik, 2017. The broader question is whether the knowledge of the L2 is the product of domain-general learning procedures (Wulff and Ellis, 2018) or whether we must invoke domain-specific algorithms. Our position here is that a domain-general hypothesis-testing model would have difficulty arriving at the hypothesis 'the WH word is licensed to remain *in situ* because the pitch is compressed in the syntactic WH domain'; it must be domain-specific.

b The role of input. While we take a modular, rationalist epistemological stance, we do not deny that the environmental input plays a critical role in L2 learning; for discussion of these issues, see Lidz and Gagliardi, 2015; Yang, 2017; Yang et al., 2017. What is potentially interesting about the current study is that the participants are classroom learners of Japanese as a Foreign Language. These learners have not received very much Japanese input (compared to naturalistic learners), and the structures in question (i.e. prosodic contiguity) have not been taught explicitly in class at all. Thus, if they acquire targetlike phonology and syntax while studying a foreign language at university, to us it seems highly unlikely that this learning is the result of a simple mapping[15] of the classroom input. To us, such a result would argue for domain-specific constraints on the nature of the interlanguage grammar.

c Shallow structure. Clahsen and Felser (2006) presented a hypothesis (the Shallow Structure Hypothesis, or SSH) which attempted to account for some of the non-native-like properties of interlanguage processing and interlanguage grammars by positing shallower, less-embedded syntactic representations in the non-native speakers. Clahsen and Felser (2017) revisit the model, and note that the:

> core claim of the SSH was that unlike native speakers, even highly proficient L2 speakers tend to have problems building or manipulating abstract syntactic representations in real time and are guided more strongly than native speakers by semantic, pragmatic, probabilistic, or surface-level information.

Relevant to our project, they make explicit that the claims of the SSH are expected to hold in production tasks as well as comprehension tasks.

They also argue that since the original research in 2006 (which focused primarily on syntactic processing) there has been empirical evidence provided (Clahsen et al., 2010) which suggests that morphological processing is also affected. Thus, it would not be a stretch to infer that they would suspect that non-native speakers would also have difficulty 'building or manipulating' abstract phonological representations. And while they acknowledge that non-native speakers can have hierarchical grammatical representations, they emphasize that the L2ers underuse abstract hierarchical representations in real-time processing. We will argue that our research participants are constructing and manipulating abstract syntactic and phonological categories in their production of L2 Japanese WH questions. To the best of our knowledge, this is the first study to empirically test the SSH in the domain of phonology.[16]

d Phonetics and phonology. We are investigating (1) the phonetic property of the implementation of a pitch-level boost on WH words, and (2) the grammatical property of phonological contiguity. The logical possibilities are that our participants will demonstrate targetlike performance in both, neither, or one. It is our expectation that the production of these classroom learners will be accented phonetically, but that they have more of a chance of acquiring the target phonology.

3 Hypotheses

Before presenting the results, let us formalize our hypotheses. For all of these, we are agnostic as to the role proficiency will play but will analyse the data to see whether the self-assessed proficiency level plays a role in their production.

1. We predict that our participants will be more nativelike on the categorical phonological phenomena than the gradient phonetic elements. Specifically, we predict that they will have no prosodic phrases marked by pitch peaks between the WH word and the complementizer.
2. We predict that our participants will be non-nativelike on the gradient phonetic phenomenon of the pitch boost on the WH word compared to DP objects. This is not because we feel the property is unlearnable, but rather that our participants are classroom learners in a foreign language environment. We do not expect them to have had the necessary amount of input[17] to have acquired this phonetic property.
3. We predict that, contra the Shallow Structure Hypothesis, our participants will show evidence of abstract, recursive phonological constituents in their interlanguage which explain the properties of their Japanese L2 production.

IV Results

The initial measurements were made and recorded in Hertz. Then, in order to normalize the differences between male and female voices, we converted the Hertz scale to semitones using the formula (12*log2(Hz/100)). Furthermore, as noted earlier, the difference between L1-English participants and L1-other participants was not significant (p = 0.315). This was based on a likelihood ratio test (LRT) of two nested linear mixed models: A full model with a random subject effect and fixed effects for object, first language, and the interaction of object and first language, and a reduced model without the interaction term. When it comes to the pitch differences on WH objects vs. DP objects we will report on all participants together.

1 Pitch boost

Let us now turn to a discussion of the first component of Richards' notion of contiguity by looking at the instantiation of pitch boost by comparing sentences (in particular the bolded items) of the type shown in (23)–(25).

(23) a. Noboru-wa **piza-o**[18] mottekitandesu ka?
Noboru-TOP pizza-ACC bring- POL [+Q]
'Did Noburo bring pizza?'

b. Tarō-wa **nani-o** mottekitandesu ka?
Taro-TOP what-ACC bring- POL [+Q]
'What did Taro bring?'

(24) a. Miki-wa kino **nani**-o kai-mash-ita ka?
Miki-TOP yesterday what-ACC buy-POL-PAST [+Q]
'What did Miki buy yesterday?'

b. Ponyo-wa kino **hon**-o kai-mash-ita ka?
Ponyo-TOP yesterday book-ACC buy-POL-PAST [+Q]
'Did Ponyo buy a book yesterday?'

(25) a. Naoya-ga **nanika**-o nomiya-de nonda.
Naoya-NOM something-ACC bar-LOC drank
'Naoya drank something at the bar.'

b. Naoya-wa **nani**-o nomiya-de nonda no?
Naoya-TOP what-ACC Bar-LOC drank [+Q]
'What did Naoya drink at the bar?'

We compared (by paired t-test, since we have paired observations for the data in sentence pairs 1–3) the average pitch on DP direct objects vs. WH direct objects in these sentences across all 16 participants, shown in Table 1.[19] None of the three pairs of sentences had a significant difference in pitch.

The preceding comparisons are illustrative but let us also compare across all of the sentences. Remember that Richards would predict that the pitch of WH phrases would be

Table 1. WH vs. DP pitch levels (in semitones) by sentence pair.

Sentence pair	Object	Average pitch (SD in parentheses)	Average difference in pitch (95% configence interval)	Significance from paired *t*-test
1	DP	12.98 (6.79)	–0.80 (–2.57, 0.97)	$p = 0.349$
	WH	12.18 (7.65)		
2	DP	13.59 (6.55)	–0.04 (–1.17, 1.08)	$p = 0.935$
	WH	13.54 (7.10)		
3	DP	11.16 (6.65)	–0.90 (–2.36, 0.56)	$p = 0.207$
	WH	10.26 (6.44)		

Table 2. WH vs. DP pitch levels all sentences.

All L2 participants' DP direct objects average	12.90
All L2 participants' WH objects average	12.78

boosted compared to comparable non-WH phrases (which we call DPs for sake of contrast). This comparison is shown in Table 2.

Across all non-native Japanese-speaking participants there is no difference in the mean pitch of WH and DP words ($p = 0.801$ on a paired *t*-test). We used a linear mixed model to account for the repeated measures on the same participants, and this also revealed no significant difference between WH and DP ($p = 0.745$, LRT of full model with random subject effect and fixed object effect and reduced model without the object effect). Furthermore, after controlling for proficiency level (as a fixed effect with the categories being either Intermediate or Advanced), we do not see differences between sentence types (DP vs. WH) ($p = 0.197$, LRT of full model with random subject effect and fixed effects for sentence type, proficiency, and the interaction of sentence type and proficiency, and a reduced model without the interaction term), indicating that the intermediate participants were not behaving differently from the advanced participants.

The lack of significant difference between the two sentence types, has, we think, two interesting implications. First, that the non-native speakers were not phonetically implementing the WH pitch boost found in native speakers. Second, this reveals that we are not seeing English echo questions being produced here ('Ponyo bought WHAT?'), but rather L2 Japanese WH questions. The L2 learners are not transferring English echo-question intonation here.

For in addition to WH-moved questions ('What should Ponyo eat?'), English allows WH *in situ* questions with a particular pragmatic force. A sentence such as *Ponyo ate WHAT?* could be used to indicate shock or surprise at what Ponyo ate. It could also be used to indicate uncertainty on the part of the listener as to what was actually said. In either case, the *what* receives a boost in pitch to indicate the marked interpretation. If the L1 English speakers were transferring the intonation from these structures then we might have expected a pitch boost on the Japanese *in situ* utterances. However, we see no

Table 3. Pitch levels (in semitones) between WH and [+Q].

Participant number	nani-o WH	nomiya-de	nonda	no [+Q]
S1	5.94	0.511	1.33	5.82
S15	20.5	15.3	15.3	17.83

significant boost. We have no reason to doubt that our participants are producing Japanese WH questions in our production task.

2 Prosodic structure

Let us turn to the second component of contiguity theory: prosodic contiguity.[20] Table 3 shows the data from two of the advanced speakers for the sentence given in (20b). The boxed cells clearly indicate the level pitch between the WH Phrase (*nanio*) and the question particle (*no*).

This almost-completely level pitch contour between the WH word and [+Q] is clearly consistent with Richards' hypothesis in that there is no prosodic boundary (as would be indicated by a pitch rise) in what he calls the 'WH domain'. For these speakers, we posit the structures (from Richards, 2010) given in (26).

(26) $[_{DP}$ WH $[_{DP}$ $[_{VP}]]$ $_{C}$
(MinP) (MinP)
(MinP) ← WH domain

The advanced participants shown in Table 3 clearly show a nativelike prosodic pattern insofar as they maintain a level pitch contour between the WH word and the question particle, with higher pitch on the WH word and on the question particle. This particular sentence pair is used for illustration but we see the same pattern in other sentences.

In order to probe the question more thoroughly, we needed to operationalize the diagnostic of a prosodic boundary. As before, the pitch values key points of the target sentences were read off Praat. We looked at the 4 sentences given in (27).

(27) a. Naoya-wa nani-o nomiya-de nonda no?
Naoya-TOP what-ACC bar-LOC drank Q
'What did Naoya drink at the bar?'

b. Taro-wa nani-o mottekita-ndesu ka?
Taro-TOP what-ACC bought-POL Q
'What did Taro bring?'

c. Dare-ga nani-o kaimasita ka?
who-NOM what-ACC bought-POL Q
'Who bought what?'

d. Miki-wa kino nani-o kaimasita ka?
Miki-TOP yesterday what-ACC bought-POL Q
'What did Miki buy yesterday?'

In each of these sentences, we looked at the change in pitch from Word1 to Word4 as indicated in Table 4.

Table 4. Sentential pitch patterns.

Sentence	Word1	Word2	Word3	Word4
1	nani-o	nomiya-de	nonda	no
2	nani-o	mottekita	ndesu	ka
3	dare-ga	nani-o	kaimasita	ka
4	kino	nani-o	kaimasita	ka

In order to analyse our data set as a whole, we had to merge the male and female voices which, of course, have different pitch ranges. First, we needed to determine that the variation in scores was the same in the male and female groups. To do this, we conducted an F-test of equality of variances for the two groups. Two-sample F-tests for the equality of variances were performed at the alpha = .05 level, and the hypothesis (that the variances are equal) was not rejected.[21] Given that the variances were equal, we were justified in adjusting the location of the mean of one of the groups rather than scaling the data. After converting Hertz scores to semitones, the males' pitch was on average 11 semitones lower for each sentence word compared to the females' pitch. Therefore, in order to make comparison across all participants, we added 11 to each male measurement such that the ratio of means of the females and males would be approximately 1 indicating a similar mean pitch across the two genders. As a result of this shift, any differences which are observed between the pitch of words, are less likely to be the result of differences between the pitch of male and female voices. The working hypothesis for us to test with respect to phonological contiguity, then, would be that there would not be a pitch increase from Word1 to Word2 or from Word2 to Word3. In other words, we predict that in terms of pitch Word3 $<$ Word2, Word3 $<$ Word1, and Word2 $<$ Word1.

a Linear Mixed Effects (LME) analyses. For each word pair comparison, we fit two linear mixed models to the data: A full model where speakers were random effects and the fixed effects were word, proficiency, and the word-proficiency interaction, followed by a reduced model with the proficiency terms removed. The fit of the full model is examined to see whether the intermediate participants were behaving differently from the advanced participants while an LRT was performed to probe the proficiency effect.

b Word2/Word1. The full model showed that the intermediate and advanced groups did not behave significantly differently ($p = 0.823$) and the LRT revealed that proficiency was not a significant factor ($p = 0.975$) in explaining the pitch relationship of Word2 to Word1. We are thus justified in combining the intermediate and advanced participants.

Table 5. Results of linear mixed models for the comparison of Word1 and Word2.

	Estimate	Standard error	*t*	*p*
Full model:				
(Intercept)	17.37	0.667	26.06	< 0.001
Word2 vs. Word1	–1.65	0.685	–2.41	0.016
Proficiency (advanced vs. intermediate)	–0.13	0.916	–.14	0.888
Word2: Proficiency advanced	0.21	0.941	0.22	0.823
Reduced model:				
(Intercept)	17.31	0.457	37.83	< 0.001
Word2 vs. Word1	–1.54	0.469	–3.28	0.001

Table 6. Results of linear mixed models for the comparison of Word1 and Word3.

	Estimate	Standard error	*t*	*p*
Full model:				
(Intercept)	17.37	0.574	30.26	< 0.001
Word3 vs. Word1	–3.79	0.644	–5.88	< 0.001
Proficiency (advanced vs. intermediate)	–0.13	0.789	–0.16	0.87
Word3: Proficiency advanced	0.27	0.885	0.31	0.76
Reduced model:				
(Intercept)	17.31	0.394	43.93	< 0.001
Word3 vs. Word1	–3.64	0.442	–8.24	< 0.001

Both the reduced and full models showed that the average pitch of Word2 is significantly less than the average pitch of Word1 ($p < 0.001$, $p = 0.016$, respectively). Results are given in Table 5.

c Word3/Word1. The full model showed that the intermediate and advanced groups did not behave significantly differently ($p = 0.757$) and the LRT revealed that proficiency was not a significant factor ($p = 0.953$) in explaining the pitch relationship of Word3 to Word1. We are thus justified in combining the intermediate and advanced participants. Both the reduced and full models showed that the average pitch of Word3 is significantly less than the average pitch of Word1 ($p < 0.001$). Results are given in Table 6.

d Word3/Word2. The full model showed that the intermediate and advanced groups did not behave significantly differently ($p = 0.946$) and the LRT revealed that proficiency was not a significant factor ($p = 0.978$) in explaining the pitch relationship of Word3 to Word2. We are thus justified in combining the intermediate and advanced participants. Both the reduced and full models showed that the average pitch of Word3 is significantly less than the average pitch of Word2 ($p < 0.001$, $p = 0.002$, respectively). Results are given in Table 7.

Table 7. Results of linear mixed models for the comparison of Word2 and Word3.

	Estimate	Standard error	*t*	*p*
Full model:				
(Intercept)	15.72	0.542	29	< 0.001
Word3 vs. Word2	–2.14	0.687	–3.11	0.002
Proficiency (advanced vs. intermediate)	0.08	0.745	0.11	0.91
Word3: Proficiency advanced	0.06	0.944	0.068	0.95
Reduced model:				
(Intercept)	15.76	0.372	42.36	< 0.001
Word3 vs. Word2	–2.10	–.471	–4.46	< 0.001

Table 8. Native vs. non-native pitch patterns.

Sequence	Significance	Comparison
Word2 vs. Word1	$p = 0.171$	NS = NNS
Word3 vs. Word1	$*p < 0.001$	NS ≠ NNS
Word3 vs. Word2	$*p = 0.001$	NS ≠ NNS

e *Summary.* These data thus far reveal that (a) the intermediate and advanced groups are behaving the same when it comes to pitch patterns, and (b) that there is a significant decline in fundamental frequency from Word1 to Word2 and from Word2 to Word3. In other words, there are no prosodic boundaries between word One and word Four.

f Interlanguage grammars. These data suggest that the non-native speakers have acquired a grammar in which the phonological contiguity of the WH phrase and the question particle is maintained. This does have broader implications which we will return to in section 3.3 but for now we note that (a) this is not a property which could have been transferred from L1 English, and (b) that this is not a property which was taught in class. This latter point was established in the debriefing session of the participants (some of the native speaker controls were the Japanese teachers). When we went over the goals of the research (after the data collection), all participants and all teachers confirmed that this was not something covered in the textbooks or in the classroom instruction.

We will not go into statistical detail here, but we also probed to see if the non-native speakers were performing in a nativelike fashion with respect to the fundamental frequency of pitch descent. We used a 2-level proficiency variable: native or non-native. The results from the LRT comparing a full model with random effects for speaker and fixed effects for word, proficiency, and the word-proficiency interaction to a reduced model without proficiency terms are shown in Table 8. This indicates that for the pitch drop from Word1 to Word2 the non-native speakers were behaving like the native speakers but for the other sequences (while there was, on average, a decline from Word2 to Word3 and from Word1 to Word3, these differences between groups were significant meaning that the non-native speakers were not implementing the decrease in fundamental frequency in a nativelike way.

It is worth noting that these two points are not incompatible: (1) the non-native speakers had significant categorical changes between all three target words, (2) the changes were not the same as the changes native speakers make. The L2ers have acquired Japanese prosodic phrasing which matches the Japanese syntax. But they are not nativelike in the gradient pitch patterns implementing these phrases. They have accents.

Let us consider this in the broader context of assessing second language speech. There is a long tradition (often couched in the Critical Period Hypothesis) to see if non-native speakers can perform a particular task (e.g. Voice Onset Time, vowel contrast) with the range of native speakers. It is clear that age of acquisition correlates highly with accentedness of L2 speech. However, as Munro (2008: 194) notes, 'native pronunciation in the L2 is not only uncommon but unnecessary.' The nativelikeness measurement bar is a high bar, and not the only bar. As Munro and Derwing (1995) have shown, heavily accented speech can be perfectly intelligible. So, our L2ers have acquired a targetlike phonology but have not acquired nativelike phonetics.

This is reminiscent of some of the work done in the acquisition of L2 stress (Archibald, 1993; Kijak, 2009). Consider what an L2 stress error could look like. There are two possibilities. The first is that the L2er places the stress on the wrong syllable. For example, the word *aróma* being pronounced *ároma*. This is a phonological error. The second possibility is that the L2er places the stress on the correct syllable as a result of targetlike metrical settings but implements the stress prominence in a non-nativelike way. So, the L2er might have higher pitch on the second syllable of *aróma* but not have reduced vowels in the first and third syllable thus pronouncing the word [aróma] instead of [ərómə]. This would be a phonetic error. What we are seeing in the performance of our participants is targetlike phonology which achieves prosodic contiguity, but accented phonetic implementation.

3 Effects of instruction / UG-constrained interlanguage

These results suggest that the L2 learners are able to acquire the target grammar insofar as they have no pitch peaks within the WH domain. The interlanguage grammar licenses WH *in situ* by virtue of prosodic contiguity. We argue that the knowledge cannot be attributed to an effect of instruction but is rather the consequence of the interlanguage grammar being constrained by prosodic contiguity.

One reviewer has suggested that this targetlike performance is the result not of universal constraints on grammatical systems but rather the 'tacit acquisition' of Japanese phonology based on instructor input. While we would agree that the knowledge acquired is tacit in the sense that it is below conscious radar, we would not agree with the position that it arises from an implicit analysis of the sentence intonation. Crucially, there are some properties of Japanese speech that the L2 learners are not imitating correctly (e.g. the pitch boost on WH words). Our position explains why some properties are acquired while others are not. The 'tacit acquisition' argument cannot make such a distinction.

Furthermore, we would argue that this aspect of the Contiguity Principle could not have transferred from L1 English. Remember that English invokes the movement strategy to achieve linear contiguity not the strategy to achieve phonological contiguity. As the interlanguage grammars evidence phonological contiguity, the question is: What led the learners to adopt this strategy? Our position is that the learners do not even consider

hypotheses that are not found in other natural languages; phonological contiguity is an option provided in the grammar-making toolbox that is UG.

4 Interfaces

Sorace (2011) argues that interfaces between syntax and pragmatics show optionality or indeterminate acquisition. Furthermore, she suggests that the syntax–lexicon interface shows optionality when involving encyclopedic knowledge or extralinguistic factors (external) but not with semantic roles (internal). Sorace (2011: 9) summarizes by saying:

> There is sufficient evidence for important developmental differences between linguistic structures that require conditions of a formal nature within the grammar, and structures that require the integration of contextual factors.

The structures under investigation here seem well-suited to contributing to the literature on the Interface Hypothesis in two ways. First of all, this is an area where the structures in question are narrowly grammatical (and do not involve pragmatic or other external factors). Second, in probing what actually causes the residual optionality observed in the L2 learners for structures involving external interfaces, Sorace et al. (2009) and Serratrice et al. (2009) argue that it is restricted input that drives the incomplete acquisition. The participants discussed in this study are classroom learners of Japanese who have spent no extended periods in Japan. Their exposure to Japanese has not been extensive. And yet, the participants have acquired the target structure of the phonological phrasing which licenses WH *in situ*. These results then, confirm the prediction of the interface hypothesis that these formal properties are, indeed, acquirable, and yet raise questions as to whether the acquisition is, in fact, input-driven. Rather, it seems to suggest more that the L2-learner grammars are governed by universal properties (such as contiguity) which are not directly read off the input.

5 Phonetics and phonology

The empirical results of our study demonstrate a difference in ultimate attainment in the domains of phonetics and phonology. With respect to phonetics, the L2 learners were not implementing a phonetic boost on WH elements compared to DP elements (as has been reported in the literature (Hirotani, 2005; Ishihara, 2003; cited in Richards, 2010). However, they were encoding a representational system in which the grammars were constrained by universal principles; principles which cannot be directly inferred from the primary linguistic data, and which are not addressed explicitly via instruction.

Given such poverty of the stimulus effects (see Özcelik and Sprouse, 2016) in SLA, this has certain architectural implications which, to our minds, justify a well-defined phonetics/phonology demarcation. It is uncontroversial to note that phonology interfaces with other motoric and grammatical systems. The interface with morphology is crucial (Distributed Morphology: Embick, 2010; Prosodic Morphology: McCarthy and Prince, 1996; Zimmermann, 2017; Lexical Phonology: Mohanan, 1986; Rubach, 2008). The interface with syntax is critical (Selkirk, 2011). However, there is much work which

blurs the line between phonetics and phonology (Boersma and Pater, 2016; Flemming, 2017). This view is in opposition to the stance of phonology as cognition (Burton Roberts et al., 1999; Reiss and Hale, 2008). The roots of such a computational model of phonology may well stem from the tenets of minimalist architecture which view spell-out to Phonetic Form (PF) as largely a motoric interface. Even the work on the evolution of the language faculty (Berwick and Chomsky, 2016) views the central development to be Merge but still views phonology as largely a production device. But Richards (2010, 2016) reminds us that there is an earlier interaction between phonology and syntax in which phonological properties may license syntactic structures. Drawing on work such as Selkirk (2011) and Elfner (2015), we see evidence for syntactic constituents being mapped one-to-one onto prosodic domains (contra Nespor and Vogel, 1986). Thus phonological domains are also fundamentally recursive (see also Wagner, 2010).

Thus, we can view L2 phonology as cognition and recognize that a phonetics/phonology split actually helps to account for the knowledge and behaviour of our participants. They are more successful at the categorial phonological phenomena (i.e. groupings of prosodic phrases) than they are at gradient phonetic phenomena (i.e. pitch level of a particular syntactic category). This is consistent with aspects of the literature on L2 phonology and nativelikeness. It is well-documented that it is uncommon for adult second language learners to fall within the range of native-speaker performance when it comes to gradient phonetic phenomena such as Voice Onset Time or vowel quality (Abrahamsson and Hyltenstam, 2009; Munro et al., 1996). We are not proposing a causal connection here (i.e. that these new contrasts are difficult because they are gradient). Indeed, it is well-established that some gradient contrasts are easier to acquire than others (Bohn, 1995; Polka and Bohn, 2011). What we are suggesting is that one way of interpreting the differential learner performance is that they have acquired the targetlike phonology but not target like phonetic implementation of Japanese WH questions.

6 Return to learnability

As Yang (2017) reminds us, language acquisition is a fundamentally psychological activity, and our acquisition models must take into account both universal constraints on the system, and the type of input available to the learner.

We therefore note that the learnability of the two phonetic and phonological phenomena is quite different. Phonological representations are constrained by UG (Match Theory; Contiguity). We can conceptualize it along the following lines:

> IF WH remains *in situ* THEN no prosodic boundaries between WH and [Q] particle.

The learners are unable to set up grammars which violate UG properties. This is consistent with other L2 UG properties (Archibald and Yousefi, 2018; Özcelik and Sprouse, 2016).[22]

Language-specific phonetic implementation,[23] on the other hand, must be picked up only from the environment, and, in this case at least, the patterns are much subtler, and potentially masked. The L2 learner must note the fundamental frequency from a range of speakers (men, women, children) on different syntactic categories (WH and non-WH

DPs), keep track of the effects of intonation, and pragmatics factors such as presupposition and focus, and from this determine that WH words should have higher pitch than DP objects. At the very least, we would expect it to take longer for the L2 learner to acquire this property of Japanese.

V Conclusions

In this article we have looked at some of the structural properties of L2 Japanese WH questions. In Japanese the WH words are licensed to remain *in situ* by virtue of the prosodic contiguity (Richards, 2016) properties of the phrases which have no prosodic boundaries between the WH word and the question particle. In a rehearsed-reading, sentence production task, we looked to see whether non-native speakers of Japanese who are learning the L2 in university classes (thus with relatively limited input compared to naturalistic learners) are able to acquire grammars which are constrained by such universal properties as Match Theory (Elfner, 2015; Selkirk, 2011) and Contiguity Theory (Richards, 2010, 2016). While LME analyses of the pitch contours reveal that the L2ers have not acquired the phonetic implementation of the documented pitch boost on WH words compared to non-WH DPs (i.e. they have foreign accents), our data show that the participants have acquired the pitch compression patterns indicative of having no prosodic phrases intervening between the WH word and the question particle.

The classroom learners were given explicit instruction as to the word order of Japanese WH questions. Their classroom environment appears to have been sufficient for them to acquire targetlike syntactic structure for this word order. This syntactic knowledge, in turn, generates a targetlike prosodic structure.

Thus, we argue, that the data are supportive of the position that interlanguage grammars are constrained by universal grammatical principles of Contiguity and Match theory which govern the prosodic contiguity of WH phrase licensing. We also presented these results as being counter to the Shallow Structure Hypothesis (Clahsen and Felser, 2017).

Acknowledgements

We would like to acknowledge feedback we have received from three anonymous *Second Language Research* reviewers and the Associate Editor, all of whom made suggestions which both strengthened the analysis and made it accessible to a wider readership. We also received feedback from audiences at the Canadian Linguistic Association in Calgary, and GASLA in Southampton on preliminary versions of this article. Thanks to Heather Goad and Joey Windsor for their comments at those conferences. Tess Nolan was invaluable in providing phonetic support. We would also like to thank Mary Lesperance for statistical advice.

Declaration of Conflicting Interest

The authors declare that there is no conflict of interest.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

John Archibald https://orcid.org/0000-0001-8500-0162

Notes

1. The particle *-ka* is used when there is a politeness marker in the sentence (see Miyagawa, 2012).
2. Merge is the operation that combines categories into hierarchically-structured units. For example, taking [the] and [sandwich] and building the new constituent [the sandwich]. Then taking [ate] and [the sandwich] to build [ate the sandwich].
3. A reviewer points out that scrambling of the WH phrase to the position adjacent to the [+Q] feature should also improve prosodic contiguity. We acknowledge that there are many technical aspects of Richards' model that we are not addressing in this article. Space does not permit us to explore the issues that scrambling raises.
4. These preferences are, in fact, modelled as violable constraints, the details of which are not crucial to our argument. In some languages, for example, single-word DPs behave as full phonological phrases.
5. We use the sparser Bare Phrase Structure trees to highlight the isomorphic mapping of heads and phrases.
6. A reviewer notes that we have drawn prosodic trees with phonological recursion. Following van der Hulst (2010), Féry (2010), we also assume that phonological structure is recursive.
7. A reviewer notes that it would be interesting to see data from yes–no questions that have a *nanika* DP for comparison. We certainly agree but will have to leave this for a follow-up study. Richards (2010: 147) shows the pitch track of one yes–no question in Fukuoa Japanese (from Smith, 2005), which shows a pitch rise between the DP and the end of the sentence.
8. A reviewer suggests that this pitch compression could be the result of pragmatic effects. We acknowledge that information structure can affect intonation; however, for these data, our participants were told to assume a neutral discourse context.
9. A reviewer wonders if the complementizer might better be analysed as being in a higher prosodic phrase. We are following Richards (2010: 145) where he defines the WH domain as [C WH], 'in which C and WH are not separated by prosodic boundaries'.
10. Appendix 2 contains a summary of the participant profiles.
11. There would have been considerable variation in the number of classroom hours of instruction the participants received. Not all of them were Japanese majors. Roughly, they would have taken 4–6 courses in Japanese language with each course having approximately 39 hours of class time.
12. These sentences were vetted (and the orthography confirmed) by a linguistics graduate student who is a native speaker of Japanese. A reviewer also points out that since the stimuli have both Roman and Japanese orthography, we cannot tease apart the different reading strategies. We agree, but our goal was to get the best, most natural production that each participant was capable of, and for that reason decide to include both orthographies.
13. The morpheme-by-morpheme gloss was not part of their reading text. Both Roman and Japanese characters were included, however.
14. Our native speaker consultant gives the following pitch accent patterns on the key words: *nanio* (HLL), *nanikao* (HLLL), *pizao* (HLL), *kino* (LHH).
15. A reviewer raises the question of what a 'simple' mapping is, or what is 'easily' read off the input; for an in depth discussion of these issues, see Carroll, 2001. In brief, we would assume that phonetic properties such as nasality or stridency are robust phonetic properties easily made use of by second language learners. In contrast, the patterns of pitch variation

in a sentence has many underlying correlates such as pitch accent, sentence type (question or statement), or information structure. For a learner to notice that there were no pitch rises between two particular syntactic categories in certain sentence types is, in our view, something that would require much more complex analysis.

16. A reviewer cautions us against over-interpreting our data with respect to the Shallow Structure Hypothesis. We feel the data are relevant for two reasons: (1) While it is logically possible that a researcher could assume shallow morphological and syntactic structure but complex phonological structure, we know of no one who takes this stance. (2) Given the nature of the syntax–phonology interface, our data provide an argument that the complex phonology is a diagnostic of complex syntax, and this is also a counter-argument to the Shallow Structure Hypothesis.
17. A reviewer wonders about the quantification of an input threshold. While we feel it is beyond the scope of this article to explore how much input might lead to the acquisition of these phonetic properties, we would note that the heritage learner literature (e.g. Kupisch and Rothman, 2016) is exploring these issues fruitfully.
18. A reviewer has noted that *piza* is a loanword and that this may raise issues of loanword phonology. This is an important point, and we are aware of the rich literature on the nature of the bilingual lexicon. However, this was the only loanword in our task, and we will have no reason to believe it lead to exceptional pitch results.
19. Our research design was not created to see if our non-native participants were behaving in a nativelike fashion. The group sizes were not equal. Furthermore, it turned out that we had dialect variation in our native speaker group; variation which possibly affects the pitch boost phenomenon. For this reason, we are not reporting on the native/non-native comparison data. We accept the claim in the literature that native speakers utilize the pitch boost on the WH phrase. Our concern is whether our L2ers treat the WH phrases differently from the DPs.
20. For this analysis, we had an additional participant, thus $n = 17$.
21. All of the variances fell within the range of 0.2958693 and 4.5960944 for the critical value limits of the F distribution.
22. A reviewer notes that in contrast to the wealth of literature addressing UG principles in L2 syntax, there has been much less attention to UG in L2 phonology. Perhaps this article will trigger future research by others.
23. Motoric skills which are not governed by universal properties of the language faculty.

References

Abrahamsson N and Hyltenstam K (2009) Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning* 59: 249–306.

Archibald J and Yousefi M (2018) The redeployment of marked L1 Persian codas in the acquisition of marked L2 English onsets: Redeployment as a transition theory. Unpublished paper presented at the Conference on Central Asian Language and Linguistics, Indiana University, Bloomington, IN, USA.

Archibald J (1998) *Second language phonology*. Amsterdam: John Benjamins.

Archibald J (1993) *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht: Kluwer.

Bates D, Maechler M, Bolker B, and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Berwick R and Chomksy N (2016) *Why only us?* Cambridge, MA: MIT Press.

Bley-Vroman R (2009) The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition* 31: 175–98.

Boersma P and Pater J (2016) Convergence properties of a gradual learning algorithm for Harmonic Grammar. In: McCarthy JJ and Pater J (eds) *Harmonic grammar and harmonic serialism*. Sheffield: Equinox, pp. 389–434.

Boersma P and Weenink D (2019) *Praat: Doing phonetics by computer: Version 6.0.55*. Available at: http://www.praat.org (December 2019).

Bohn O-S (1995) Cross language speech perception in adults: First language transfer doesn't tell it all. In: Strange W (ed) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, pp. 279–304.

Burton Roberts N, Carr P, and Docherty G (1999) *Phonological knowledge: Conceptual and empirical issues*. Oxford: Oxford University Press.

Cabrelli Amaro J (2017) The role of prosodic structure in the acquisition of L2 Spanish stop lenition. *Second Language Research* 33: 233–69.

Carroll S (2001) *Input and Evidence: The Raw Material of Second Language Acquisition*. Amsterdam: John Benjamins.

Chomsky N and Halle M (1968) *The Sound Pattern of English*. New York: Harper and Row.

Clahsen H and Felser C (2006) Grammatical processing in language learners. *Applied Psycholinguistics* 27: 3–42.

Clahsen H and Felser C (2017) Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition* 40: 693–706.

Clahsen H, Felser C, Neubauer K, Sato M, and Silva R (2010) Morphological structure in native and nonnative language processing. *Language Learning* 60: 21–43.

Deguchi M and Kitagawa Y (2002) Prosody and *wh*-questions. In: Hirotani M (ed) *Proceedings of NELS 32*. Amherst, MA: GLSA, pp. 73–92.

Elfner E (2015) Recursion in prosodic phrasing: Evidence from Connemara Irish. *Natural Language and Linguistic Theory* 33: 1169–1208.

Embick D (2010) *Localism versus globalism in phonology and morphology*. Cambridge, MA: MIT Press.

Féry C (2010) Recursion in prosodic structure. *Phonological Studies (Kaitakusha, Tokyo)* 13: 51–60.

Flemming E (2017) Dispersion theory and phonology. In: Aronoff M (ed) *The Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press.

Fodor J (2002) Prosodic disambiguation in silent reading. In: Hirotani M (ed) *Proceedings of the North East Linguistics Society* 32. Amherst, MA: GSLA, pp. 112–32.

Franceschina F (2001) Morphological or syntactic deficits in near-native speakers?: An assessment of some current proposals. *Second Language Research* 17: 213–47.

Goad H and White L (2006) Ultimate attainment in interlanguage grammars: A prosodic approach. *Second Language Research* 22: 243–68.

Hirotani M (2005) Prosody and LF Interpretation: Processing Japanese wh-questions. Unpublished doctoral dissertation, University of Massachusetts. Amherst, AM, USA.

Ishihara S (2003) Intonation and interface conditions. Unpublished doctoral dissertation, MIT, Cambridge, MA, USA.

Juffs A and Harrington MW (1995) Parsing effects in L2 sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition* 17: 483–516.

Kijak A (2009) *How stressful is L2 stress?* Utrecht: LOT.

Kim E, Montrul S, and Yoon J (2015) The on-line processing of binding principles in second language acquisition: Evidence from eye tracking. *Applied Psycholinguistics* 36: 1317–74.

Kubozono H (2006) Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua* 116: 1140–70.

Kupisch T and Rothman J (2016) Terminology matters!: Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism* 22: 564–82.

Levelt WJM (1989) *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lidz J and Gagliardi A (2015) How nature meets nurture: Universal Grammar and statistical learning. *Annual Review of Linguistics* 1: 333–52.

McCarthy J and Prince A (1996) Prosodic morphology. In: Goldsmith J (ed) *The handbook of phonological theory*. Oxford: Blackwell, pp. 318–66.

Miyagawa S (2012) Agreements that occur mainly in the main clause. In: Aelbrecht L, Haegeman L, and Nye R (eds) *Main clause phenomena: New horizons*. Philadelphia, PA: John Benjamins, pp. 79–111.

Mohanan KP (1986) *The theory of lexical phonology*. Dordrecht: Springer.

Montrul S (2011) Multiple interfaces and incomplete acquisition. *Lingua* 121: 591–604.

Munro M (2008) Foreign accent and speech intelligibility. In: Hansen Edwards J and Zampini M (eds) *Phonology and second language acquisition*. Amsterdam: John Benjamins, pp. 193–218.

Munro MJ and Derwing TM (1995) Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45: 73–97.

Munro M, Flege J and Mackay I (1996) The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313–334.

Nespor M and Vogel I (1986) *Prosodic phonology*. Dordrecht: Foris.

Özcelik O (2017) Universal Grammar and second language phonology: Full Transfer/prevalent access in the L2 acquisition of Turkish 'stress' by English and French speakers. *Language Acquisition* 25: 231–67.

Özcelik Ö and Sprouse R (2016) Emergent knowledge of a universal phonological principle in the L2 acquisition of vowel harmony in Turkish: A 'four'-fold poverty of the stimulus in L2 acquisition. *Second Language Research* 33: 179–206.

Pliatsikas C and Marinis T (2013) Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition* 16: 167–82.

Polka L and Bohn O-S (2011) Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39: 467–478.

R Core Team (2015) *R: A language and environment for statistical computing* [software]. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org (December 2019).

Reiss C and Hale M (2008) *The phonological enterprise*. Oxford: Oxford University Press.

Richards N (2010) *Uttering trees*. Cambridge, MA: MIT Press.

Richards N (2016) *Contiguity theory*. Cambridge, MA: MIT Press.

Rubach J (2008) An overview of lexical phonology. *Language and Linguistics Compass* 2: 456–77.

Selkirk E (2011) The syntax–phonology interface. In: Goldsmith J et al. (eds) *The handbook of phonological theory*. 2nd edition. Oxford: Blackwell, pp. 435–84.

Serratrice L, Sorace A, Filiaci F, and Baldo M (2009) Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness. *Bilingualism: Language and Cognition* 12: 1–19.

Smith J (2005) On the *wh*-question intonational domain in Fukuoa Japanese: Some implications for the syntax–prosody interface. In: Kawahara S (ed) *UMOP 30: Papers on Prosody*. Amherst, MA: GSLA.

Sorace A and Filiaci F (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22: 339–68.

Sorace A, Serratrice L, Filiaci F, and Baldo M (2009) Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119: 460–77.
Sorace A (2011) Pinning down the concept of interface in bilingual development. *Linguistic Approaches to Bilingualism* 2: 209–16.
Sughara M (2003) Downtrends and post-FOCUS intonation in Tokyo Japanese. Unpublished doctoral dissertation, University of Massachusetts. Amherst, MA, USA.
Tyler M (2019) Simplifying MATCH WORD: Evidence from English functional categories. *Glossa* 4: 1–32.
van der Hulst H (2010) A note on recursion in phonology. In: van der Hulst H (ed) *Recursion and human language*. Berlin: Mouton de Gruyter, pp. 301–42.
Wagner M (2010) Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory* 28: 183–237.
White L (2003) *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
White L (2011) Second language acquisition at the interfaces. *Lingua* 121: 577–90.
Wulff S and Ellis NC (2018) Usage-based approaches to second language acquisition. In: Miller DT, Bayram F, Rothman J, and Serratrice L (eds) *Bilingual cognition and language: The state of the science across its subfields*. Amsterdam: John Benjamins, pp. 37–56.
Yang C (2017) Rage against the machine: Evaluation metrics in the 21st century. *Language Acquisition* 24: 100–25.
Yang C, Crain S, Berwick R, Chomsky N, and Bolhuis J (2017) The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience and Biobehavioral Reviews* 81: B: 103–19.
Zimmermann E (2017) *Morphological length and prosodically defective morphemes*. Oxford: Oxford University Press.

Appendix 1

Reading sentences

(1) Satsuki ga depāto de ojōsan no tame ni ranpu o eranda.
Satsuiki-NOM dept.store-LOC daughter-GEN sake-DAT lamp-ACC chose
サツキが、デパートでお嬢さんのためにランプを選んだ。
'Satsuki chose a lamp for her daughter at the department store.'

(2) Naoya ga nanika o nomiya de nonda.
Naoya-NOM something-ACC bar-LOC drank
ナオヤが、何かを飲み屋で飲んだ。
'Naoya drank something at the bar.'

(3) Naoya wa nani o nomiya de nonda no?
Naoya-TOP what-ACC bar-LOC drank Q
ナオヤは、何を飲み屋で飲んだの？
'What did Naoya drink at the bar?'

(4) Umi ga depāto de dare no tame ni ranpu o eranda no?
Umi-NOM dept.store-LOC who-GEN sake-DAT lamp-ACC chose Q
ウミが、デパートで誰のためにランプを選んだの？
'Who did Umi choose a lamp for at the department store?'

(5) Dare ga depāto de dare no tame ni ranpu o eranda no?
Who-NOM dept.store-LOC who-DAT sake-DAT lamp-ACC chose Q
誰が、デパートで誰のためにランプを選んだの？
'Who chose a lamp for whom at the department store?'

(6) Anata wa Haru ga nani o suki ka shitteimasu ka?
you-TOP Haru-NOM what-ACC like Q know Q
あなたは、ハルが何を好きか知っていますか？
'What do you know Haru likes?'

(7) Anata wa dare ga Tatsuo o suki ka shitteimasu ka?
you-TOP who-NOM Tatsuo-ACC like Q know Q
あなたは、誰がタツオを好きか知っていますか？
'Who do you think likes Tatsuo?'

(8) Dare ga nani o kai-masita ka?
who-NOM what-ACC bought-POL Q
誰が何を買いましたか？
'Who bought what?'

(9) Seiko wa dare ga tsukutta sushi o tabeta no?
Seiko-TOP who-NOM made sushi-ACC ate Q
セイコは、誰が作った寿司を食べたの？
'Seiko ate sushi that who made?'

(10) Shiro wa hon'ya de nani o katta no?
Shiro-TOP bookstore-LOC what-ACC bought Q
シロは、本屋で何を買ったの？
'What did Shiro buy at the bookstore?'

(11) Kumiko wa doko de zasshi o katta no?
Kumiko-TOP where-LOC magazine-ACC bought Q
クミコは、どこで雑誌を買ったの？
'Where did Kumiko buy a magazine?'

(12) Hayao ga senshū katta hon ga nakunatta
Hayao-NOM last week bought book-NOM disappeared
ハヤオが先週買った本が、無くなった。
'The book Hayao bought last week disappeared.'

(13) Masahiro ga katta hon ga itsu nakunatta no?
Masahiro-NOM bought book-NOM when disappeared Q
マサヒロが買った本は、いつ無くなったの？
'When did the book Masahiro bought disappear?'

(14) Ponyo wa kinō hon o kai-mashita ka?
Ponyo-TOP yesterday book-ACC bought-POL Q
ポニョは、昨日本を買いましたか？
'Did Ponyo buy a book yesterday?'

(15) Miki wa kinō nani o kai-mashita ka?
Miki-TOP yesterday what-ACC bought-POL Q
ミキは、昨日何を買いましたか？
'What did Miki buy yesterday?'

(16) Miki wa kinō hon o kai-mashita.
Miki-TOP yesterday book-ACC bought-POL
ミキは、昨日本を買いました。
'Miki bought a book yesterday.'

(17) Noboru wa piza o mottekitandesu ka?
Noburu-TOP pizza-ACC brought-POL Q
ノボルは、ピザを持って来たんですか？
'Did Noburu bring pizza?'

(18) Noboru wa kinō piza o kai-mashita.
Noburu-TOP yesterday pizza-ACC bought-POL
ノボルは、昨日ピザを買いました。
'Nobouru brought pizza yesterday.'

(19) Tarō wa nani o mottekitandesu ka?
Taro-TOP what-ACC bought-POL Q
タローは、何を持って来たんですか？
'What did Taro bring?'

Appendix 2. Participant profiles.

Speaker	Sex	Proficiency	Either parent speaks Japanese	Mainly instructed or natural input?	Length of residency in Japan
TI	Male	Intermediate	No	Instructed	> 1 Year
JM	Female	Intermediate	No	Instructed	< 1 year
KR	Female	Intermediate	No	Instructed	< 1 year
CS	Female	Advanced	No	Instructed	< 1 month
A	Male	Advanced	No	Natural	< 1 month
RKV	Male	Intermediate	No	Instructed	< 1 year
KSD	Female	Advanced	No	Instructed	< 1 month
ZM	Male	Advanced	No	Natural	< 1 month
TF	Female	Advanced	No	I = N	> 1 year
CJ	Male	Intermediate	No	Both	> 1 year
OW	Male	Advanced	No	I = N	< 1 year
AK	Female	Intermediate	Yes	Natural	< 1 year
JJ	Female	Intermediate	No	Natural	< 1 year
RC	Male	Intermediate	No	Natural	< 1 year
SK1	Female	Intermediate	Yes	i = N	< 1 year
NA	Female	Advanced	Yes	Natural	< 1 month
SK2	Female	Advanced	No	Natural	> 1 year