Japanese: Phonetic Overview and Suprasegmentals

Dickson Tsai

University of California, Berkeley

Linguistics 110 Term Paper

December 13, 2013

1. INTRODUCTION

1.1 The Japanese Language

Japanese is the ninth most popular language of the world, spoken by 126 million people (Nationalencyklopedin 2010). Japan is part of the Japonic language family, which also includes Ryukyuan languages (from the islands south of Japan) such as Okinawan (Lewis et al 2013). Some scholars have linked Japan to the Altaic family (e.g. Turkish, Mongolian, Korean), although it remains controversial. Japanese is the de facto national language of Japan, and is recognized as a regional language (not official) in Palau. Other speakers are dispersed among countries such as Taiwan, Korea, China, Brazil, Peru, and the United States, where 436,110 people speak Japanese at home[1], partly a result of Japan's activity during World War II, and partly from emigration. Within Japan, there are various local dialects, clustered in regions such as Osaka, Kansai, and Tokyo, although the education system and media have employed Standard Japanese that best matches the Tokyo dialect (Johnston 2007). Apart from people from the more rural parts of Japan, Japanese speakers are not monolingual – emigrants have to speak their home language and most Japanese students learn English at school, according to my language consultant.

1.2 Language Consultant

I worked with my language consultant Ena Hariyoshi, a 19-year-old female EECS student at UC Berkeley. She emigrated from Japan when she was seven years old, later residing in Michigan before arriving at UC Berkeley. She was primarily exposed to Japanese from her parents, who are both Chinese but moved to Japan in college, from Japanese school, and from her Japanese friends, all with emphasis on the Tokyo dialect. Her Chinese knowledge is limited to understanding basic words.

1.3 Overview of the Report

Unlike English, Japanese has the MORA as its most basic phonological unit[23], reflected in the orthography. The CV and V sequences, the end nasal N, and the first consonant of geminates Q, are all considered one mora. This term will be used throughout. Since Japanese is a heavily researched language, this report will provide a brief yet comprehensive phonetic inventory, highlighting sounds that are not found in English, but will mainly focus on suprasegmentals as noted in the title.

2. VOWELS

Japanese has 5 vowels in its vowel system – [i], [e], [a], [o], [ɯ] – which can become nasalized before the nasal coda or voiceless in the case for [i], [o], and [ʉ]. [a] falls between the fourth and fifth cardinal vowel, while [i] is nearly the same as the first cardinal vowel. Both [e] and [o] are lower than their respective cardinal vowels. Finally, [ɯ] is more to the front than its matching cardinal vowel with an additional complication: as demonstrated by my language consultant, even though the lips are not rounded, they are still pressed close to each other[4]. According to my language assistant, there are no diphthongs: all vowels are pronounced equally. However, some diphthongs are actually just a long vowel, such as /ei/->[eː] in *jinsei* "life"[5].

Table 1. Formant measurements from the /k/ and /s/ morpheme series.

Vowel		Example Word						CD Track
Roman.	IPA	Rom.	IPA	Definition	F1 (Hz)	F2 (Hz)	F3 (Hz)	# - Name
i	i	ki	ki	tree	388.44	2801.3	3239.2	
		shi	ɕi	death	414.26	2473.4	3100.2	
e	e̞	ke	ke̞	fur	509.80	2245.1	2797.4	
		se	se̞	height	528.84	2045.3	2799.5	
a	a	ka	ka	mosquito	838.61	1401.5	2791.7	
		sa	sa	difference	786.25	1307.2	2618.5	
o	o	ko	ko̞	child	549.44	862.35	2953.8	
		soto[a]	soto̞̬	outside	581.39	1030.6	2935.7	
u	ɯ	ku	kɯ	9 (can be [kʲɯ] too)	414.01	1536.6	2675.5	
		su[b]	sʉ̞	vinegar	426.37	1883.3	2647.5	

[a] The second [o] of "soto" is measured here because of less vowel movement
[b] My consultant's "u" sound can be considered central here, and the measurements reflect that fact (F2 is much higher than in "ku"). See Note 2 for more discussion.

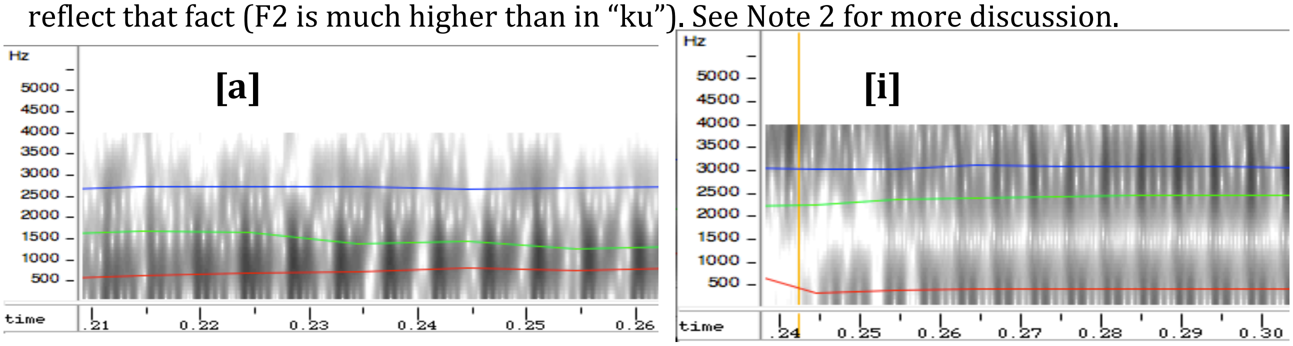


Figure 1. Plots of [a] from "sa" and [i] from "shi". The spectrogram is cut off at 4000 Hz due to background noise.

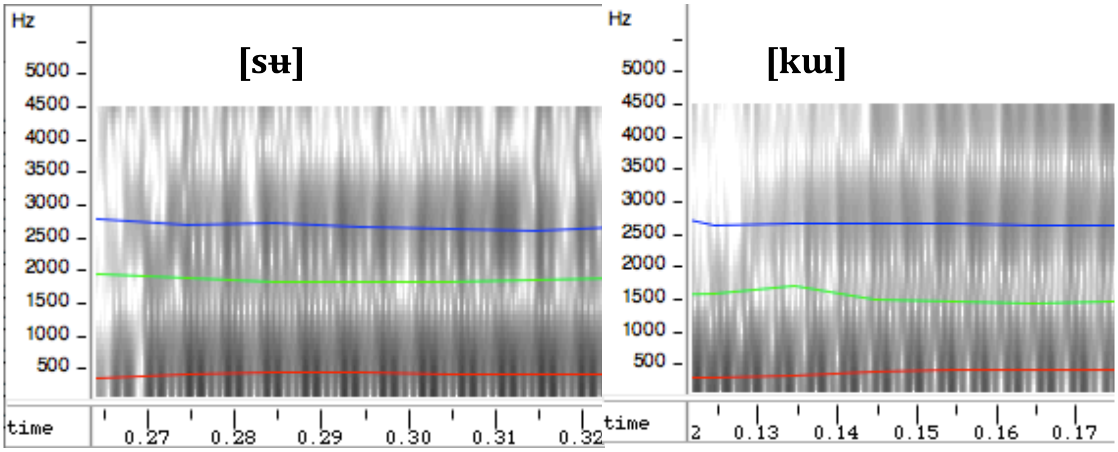


Figure 2. Plots of "u" in "su" and "ku". Notice that F2 is significantly higher in "su", which leads to the difference in transcription.

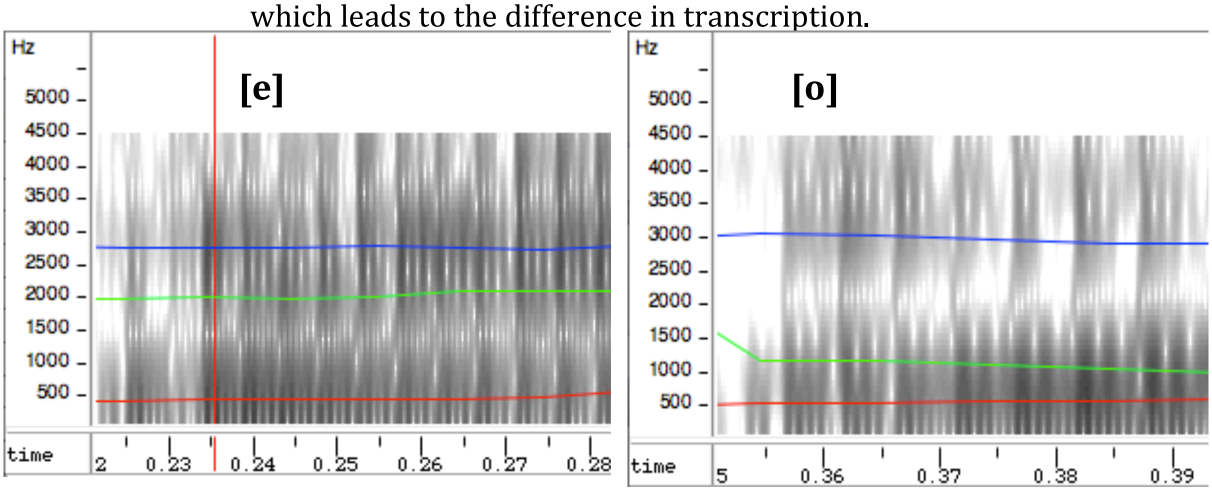


Figure 3. Plots of [e] in "se" and [o] in "soto".

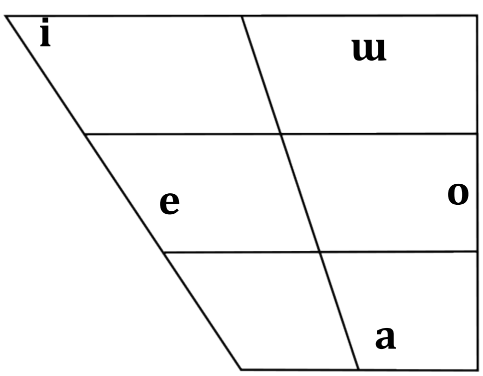


Figure 4. Japanese vowel chart based on personal estimations from the cardinal vowel system.

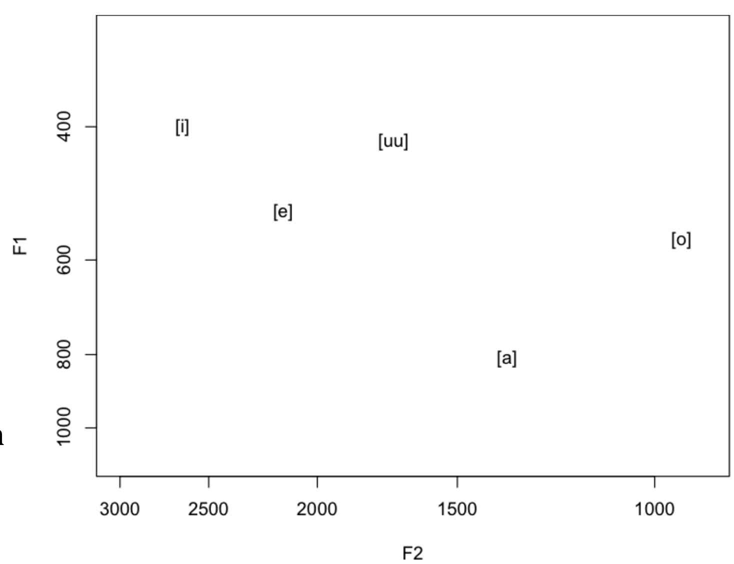
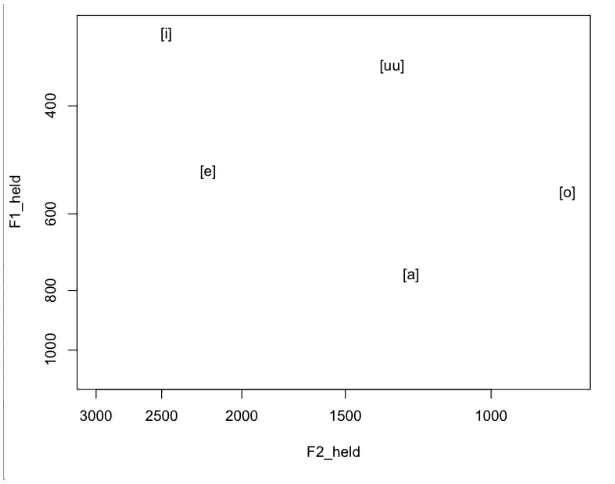


Figure 5. Log-log plot of F1 vs. F2 of the morphemes in Table 3. Each vowel's value in the plot is the average of the values given by the two morphemes.

	F1	F2	F3
i	306.16	2466.8	2741.4
e	514.05	2196.4	2857.8
a	756.06	1248.6	3036.9
o	556.21	808.74	3057.2
u	345.60	1317.0	2550.4

Figure 6. The plot with data for vowels that are purely held out. This gives more consistent measurements, and most of these single vowels are themselves morphemes.



3. CONSONANTS

	Bilabial	Alveolar	Alveolo-Palatal	Palatal	Velar	Glottal
Stop	p b	t d			k g	
Palatalized Stop	pʲ bʲ				kʲ gʲ	
Fricative	Φ ^β	s z	ɕ ʑ	ç		h
Affricate		t͡s	c͡ɕ ɟ͡ʑ			
Flap		*ɾ				
Approximant				j	ɰ	
Lateral Approximant		ɺ ɺʲ				
Nasal	m	n		*ɲ	*ŋ	
Palatalized Nasal	mʲ	nʲ			*ŋʲ	

Table 2. The consonant inventory of Japanese. [β] only appears in loanwords such as "violin". Akematsu (1997 p. 138) does not include this sound in his table. The asterisked phones do not normally appear word-initially unless conditioned by the previous word in fast speech (vs. citation). The right column indicates voicing.

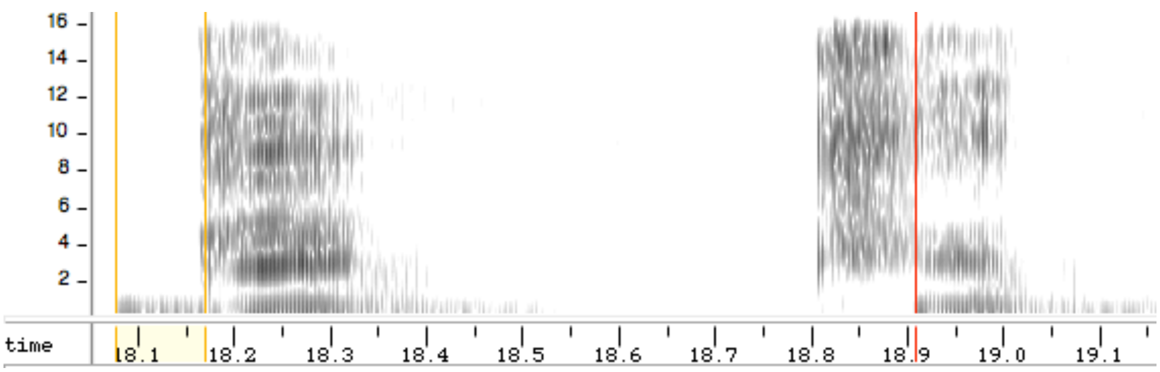
3.1 Stops

Stop	Roman.	IPA	Definition	CD Track #
p	pachipachi	['paɕi̥paɕi]	clap (onomatopoeia)	
b	ba	ba	place	
t	tate	tʰate̞	vertical	
d	udon	ɯdoŋ	wheat noodle	
k	kin	kin	gold	
g	gin	gin	silver	

Table 3. Stop Consonants

Generally, the stops in Japanese are similar to those in English, the only major difference being that the [p], [t], and [k] are only slightly aspirated. The VOT from my language assistant measured on average about 0.065 seconds for [t] and 0.094 seconds for [k] (both word-initial) past the release of the stop for the voiceless stop series. Although the second [t] in *tate* is un-aspirated in Table 3, [t^h] versus [t] is by no means contrastive. On the other hand, Figure 6 shows that my language assistant starts voicing 0.08 seconds ahead of her release for [g], although my other recordings show that there is more fluctuation, with one [g] not having any voicing at all before release.

Figure 7. Comparison of VOT in gin "silver" and kin "gold". The pre-voicing for gin are marked by double vertical lines, while the aspiration for kin is marked by one solid line. The y-axis is in kHz.



3.2 Fricatives and Affricates

Phone	Romanization	IPA	Definition
ɸ	futon	ɸɯ̥tõŋ	futon/bed
β	vaiolin	βaioɹlin	violin
s	saru	saɹɯ	monkey
z	zaru	zaɾɯ	wicker basket
ɕ	shi	ɕi	death
ʑ	ji	ʑi	letter
ç	hi	çi	son
h	ha	ha	teeth
t͡s	tsuchi	t͡sɯ̥c͡ɕi	dirt
c͡ɕ	chi	c͡ɕi	blood
ɟ͡ʑ	jinsei	ɟ͡ʑĩseː	life

Table 4. Japanese fricative and affricates. There are 8 fricatives and 3 affricates.

Apart from [s] and [z], Japanese have fricatives that sound like English but are sufficiently different for a different transcription. The [ɕ] and [ʑ] sounds, for instance, do not have lip rounding, unlike the English [ʃ] and [ʒ], which makes the sound subjectively lighter. The Chinese fricative [ɕ] is a better comparison to the Japanese sound[6]. In Figure 8, the LPC analysis of the spectral section of *shi* "death" reveals that the first formant is at 2987 Hz, which is lower than the fundamental resonance of [ʃ] (Mannell 2008). Hence, the resonating tube must be longer for *shi*, and so we transcribe it as the alveolo-palatal [ɕ]. We do not use [ç] because there are still some high-frequency noise components that resemble [s] and [ʃ].

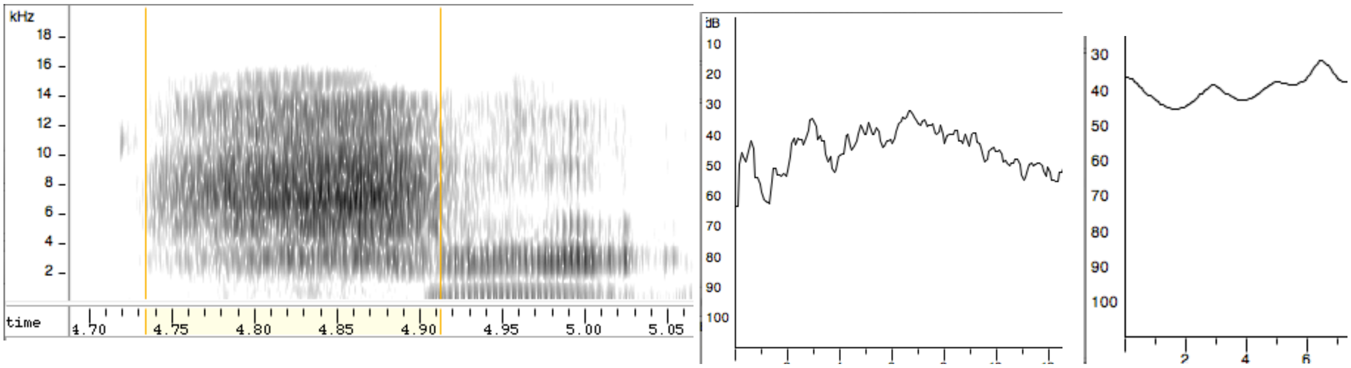


Figure 8. The spectrogram, FFT plot, and LPC plot of *shi* “death”. The vertical lines denote the slice that I used for the analyses. The first non-zero peak of the LPC is at about 2987 Hz.

The [ç], [ɸ], and [h] sounds are allophones of the same phoneme /h/ - my language consultant considers them as the same letter, and the orthography puts them in the same h series: ひ /hi/->[çi], ふ /hɯ/->[ɸɯ],ほ/ho/-> [ho][7]. Whereas the affricate [t͡s] is straightforward, the other affricates [c͡ɕ] and [ɟ͡ʑ] maintain the same qualities that the alveolo-palatal sounds have themselves. Figure 9 lends itself to the same analysis for [c͡ɕi].

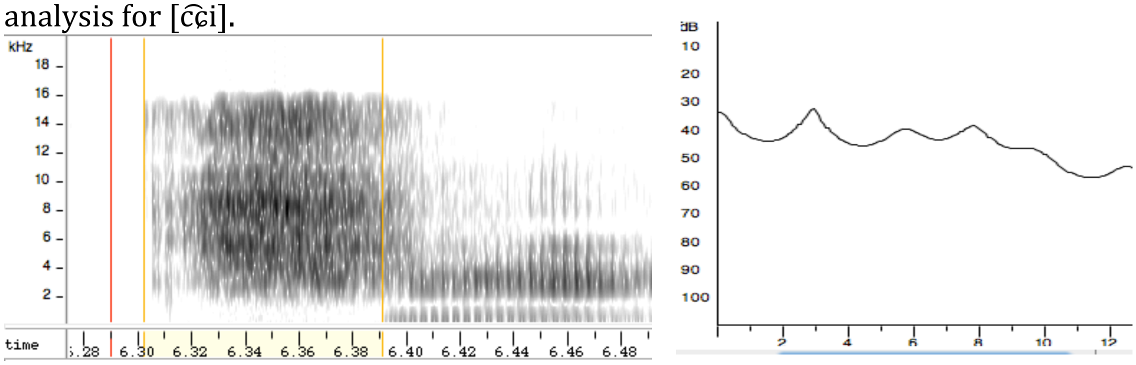


Figure 9. The spectrogram, FFT plot, and LPC plot of *chi* “blood”. The two right vertical lines denote the slice that I used for the analyses. The first non-zero peak of the LPC is at about 2936 Hz.

3.3 The Approximants + Flap

Phone	Romanization	IPA	Definition
ɺ	ruri	ɺɯɺi	lapis lazuli
ɾ	dōro	touɯɾo	road
j	ya	ja	arrow
ɰ	wakka	ɰak˺ka	hoop

The most difficult sound I had to imitate with my language consultant was the [ɺ] sound, which is right in between the [ɹ] and [l] in English. To see why, we can compare the FFT/LPC plot of [ɺ] to the FFT/LPC plots that Robert Mannell has on [ɹ] and [l]. His [ɹ] plot “has a spectrum that is strongly weighted to the low frequency region” and has a “low frequency feature with a peak at about 200 Hz and a slope down to the noise floor" (Mannell 2008). Meanwhile, [l] keeps its formants peaks but with a higher F1 and F2 than F2 and F3 (Mannell 2008). In Figure 10, our FFT/LPC plot has characteristics of both [ɹ] and [l]: first, the slope is sharply downward with the highest peak at the beginning, which is characteristic of the [ɹ]. However, one can see four distinct formant peaks in the FFT plot, which is characteristic of the [l]. While my language consultant also has trouble reconciling the Japanese [ɺ] with the English [ɹ] and [l], she notes that there is no contrast among the pair in Japanese (Personal communication with consultant).

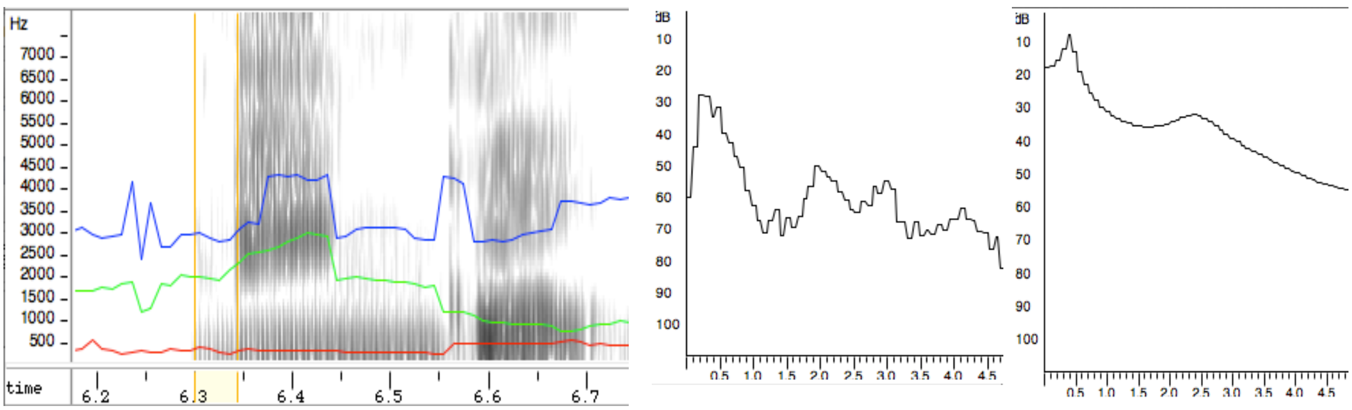


Figure 10. The spectrogram, FFT plot, and LPC plot of *ringo* "apple". The two right vertical lines denote the slice that I used, and the formant lines denote F1, F2, and F3. The spectrogram is cut off at 8000 kHz, and the FFT/LPC plots' x-axis is in kHz, with an initial peak of 384 Hz.

The Japanese [j] is similar to the English [j], and the [ɰ] sound is basically the counterpart of [w] in English, except that it is unrounded, as with the vowel.

3.4 Palatalization

Palatalized	Romanization	IPA	Definition
pʲ*	happyaku	[hap˺pʲakɯ]	800
bʲ	byooin	[bʲo:iŋ]	hospital
kʲ	Tōkyō-do	[tʰo:kʲio:do]	City of Tokyo
gʲ	gyu	[gʲɯ]	beef
ç	hyaku	[çakɯ]	100
ɾʲ	ryu	[ɾʲɯ]	dragon
ɺʲ	ringo	[ɺʲiŋgo]	apple
mʲ	myaku	[mʲakɯ]	pulse, connection
ɲ	nyanko	[ɲãŋko]	baby speak for cat
~~ŋ~~ʲ^ gʲ	atsui gyoza	[atsɯigʲoza]	hot dumpling

- *pya* is not commonly encountered word-initially, although my language consultant asserts that this consonant exists in the language, usually in newly-coined words.
- ^ I included this example because the literature has examples of /g/->[ŋ] and /g^j/->[ŋ^j] in connected speech, but my language consultant does not have this nasalization.

All of the obstruents, in addition to [ɹ] and [ɾ], can be palatalized, which means that there is a secondary articulation very similar in place to where [j] or [i] would be pronounced near the end of the hard palate. We have the spectrogram, formant plot, and power plot of the minimal pair *gyu* "cow" and *gu* "ingredients" in Figure 6. First, we notice that this secondary articulation happens in the same motion as the primary articulation because there is only one peak in the power plot for *gyu*, so *gy* is not a consonant cluster [gj] but rather a single consonant [g^j]. Next, note that the F2 starting from the burst of *gyu* is 2470 Hz compared to the F2 of *gu* at 2017 Hz. Similarly, [i] is farther front in the vowel space, which implies a higher F2, so there is a correlation between the palatalization and the frontness of the vowel [i]. Lastly, consider the difference in intensity in the stop consonant release: 58 dB for *gyu* and 48 dB for *gu*. The palatalization will create an additional obstruction in the oral cavity, which means that more pressure may build up within the area between the tongue body and velum. The speaker will then need to have greater pressure in the lungs to force air through the buildup, thereby increasing the initial intensity. The same phenomenon is not as prominent in *kyu* "nine" because no voicing is involved.

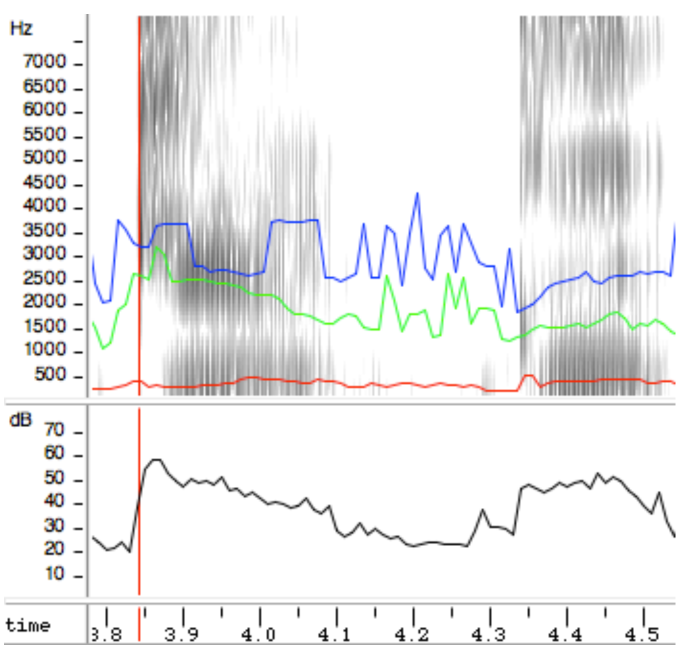


Figure 11. The spectrogram and power plot of plot of *gyu* "beef" vs *gu* "ingredients". Note the slope of the F2 line, which starts high and ends at where the F2 for [ɯ] is.

3.5 The Nasal Consonant

In Japanese, the word-initial nasal in citation speech is always [n] or [n^j]. The nasal is not always alveolar in other situations, however. Please see Section 6 – Introductory Phonology for more discussion and examples.

3.6 Geminate Consonants

Just as vowel length leads to contrasts, so do geminate consonants – for example, *tate* "vertical" and *tatte* "stand up". My language assistant considers the extra consonant a pause, although from observing her mouth as she speaks the two words, the tongue is already in position for the consonant, and so *tatte* sould be transcribed with the unreleased diacritic: [tat˺te].

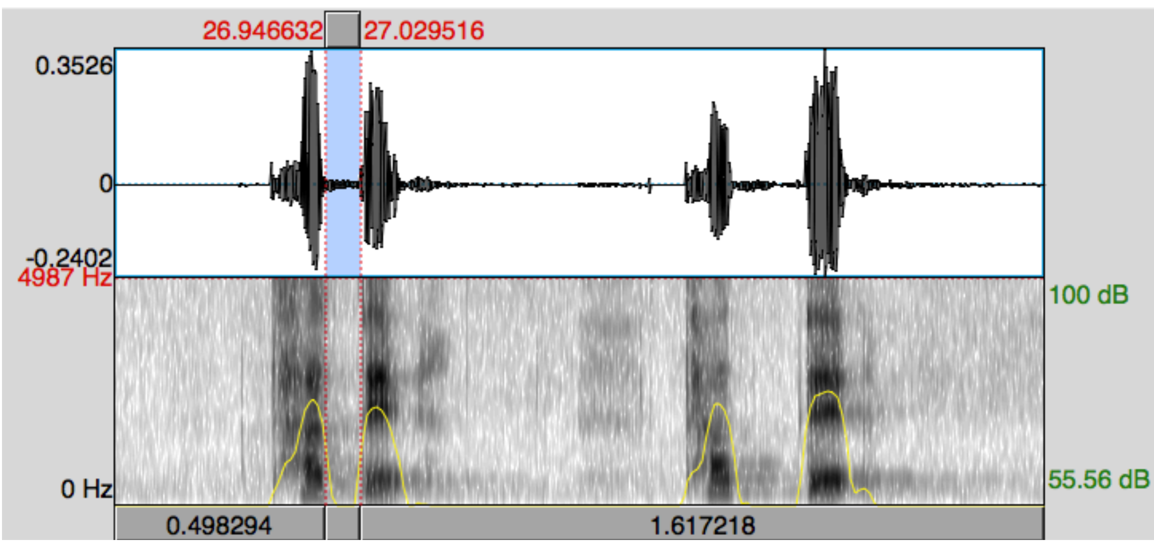


Figure 12. The waveform and spectrogram of *tate* "vertical" together with *tatte* "stand up". The silent period in *tate* is 0.083 seconds while the silent period in *tatte* is 0.1733 seconds, a ratio of 1:2.088. *tate* is a 2-mora word while *tatte* is a 3-mora word, with the first *t* being its own mora.

4. INTONATION/SPEECH MELODY

4.1 Moraic Level: Effect of Japanese Pitch Accent on the Mora

At first glance, Japanese and English may seem to share a similar stress system[8], but the manner in which both languages achieve these word-to-word distinctions is very different. In English, stress can be manifested by pushing more air through the lungs, lengthening the vowel, or raising the pitch, but the most reliable way to detect stress is to see if that syllable affects the intonation contour of the phrase (Ladefoged and Johnson 2011; Hasegawa 1995). On the other hand, Japanese stress is primarily expressed with a higher pitch. Consider Figure 13 above, a recording of

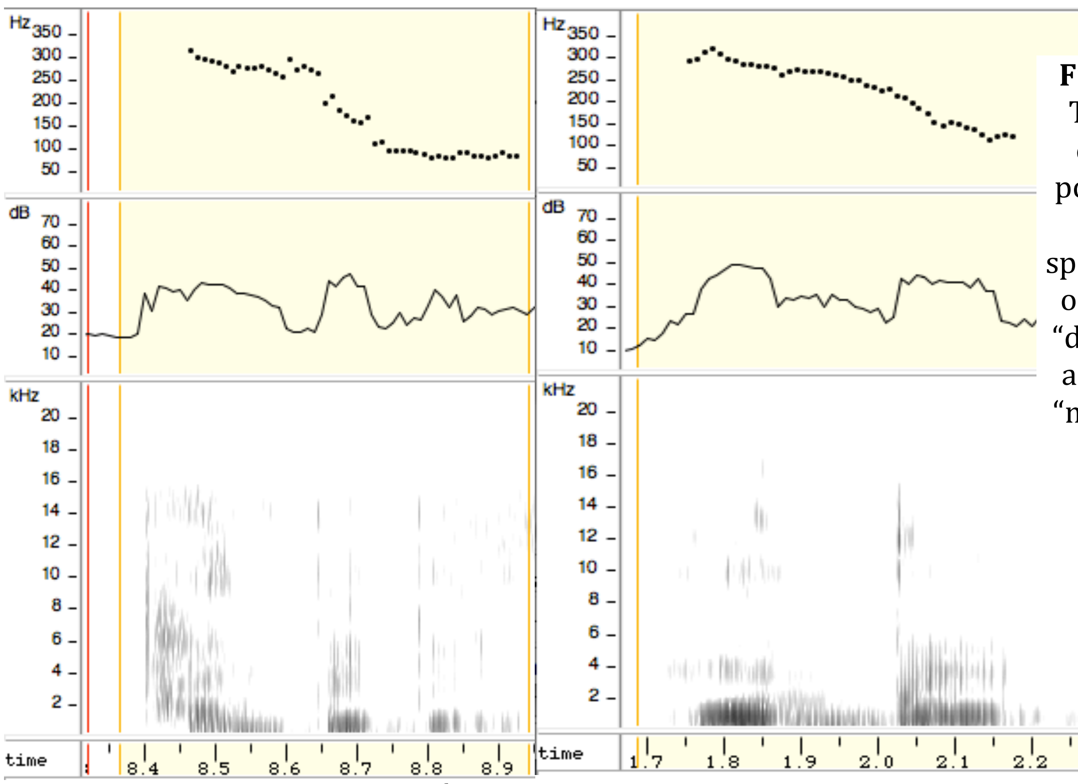


Figure 13. The pitch contour, power plot, and spectrogram of *tanpopo* “dandelion” and *hondo* “mainland”.

the word *tanpopo* “dandelion” [t^hãmpopo] and the word *hondo* “mainland” [ˈhõndo]. The pitch contour is clearly higher for *tanpopo*, at around 300 Hz for the first syllable before downstepping for the second and third syllables. Meanwhile, the intensity plot actually shows an increase from the first to the second syllable. The plots from *hondo* agree – a significant difference in pitch, but not of intensity. Japanese stress can generally apply to any syllable, and more than one syllable may have a higher pitch. Consider *nihongo* “Japanese language” [ɲiˈhõŋˌgo], where it is difficult to assign "primary" stress because both the second and third syllable have a higher pitch. However, there must be a difference in pitch for the accent to be perceived, and in *nihongo*, the first and second morae have different pitches[9].

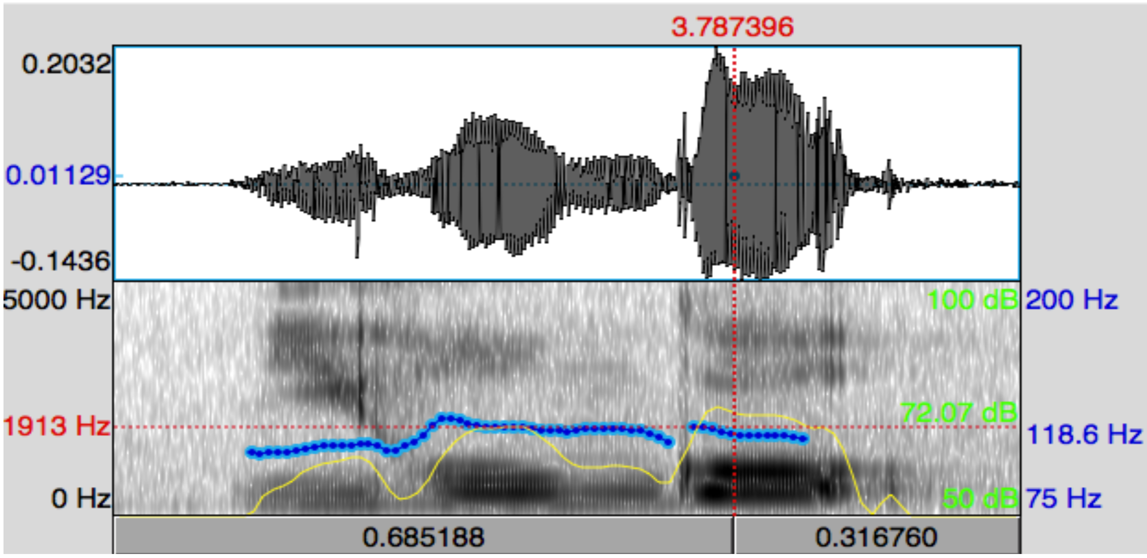
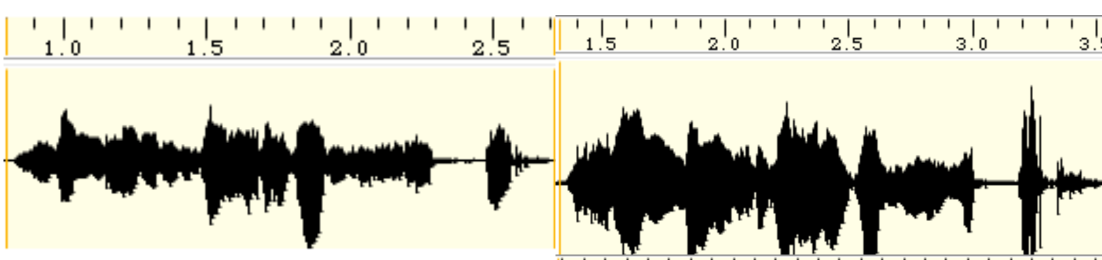


Figure 14. The waveform and spectrogram of *nihongo* "Japanese language". The thick dotted line is the pitch tracing of the word. The highest pitch in *ni* is 109.6 Hz, while the highest pitch in *hon* and *go* are initially above 118 Hz.

While English has a stress contrast on every polysyllabic word, Japanese has contrast only on a certain number of words, such as *háshi* "chopsticks" and *hashí* "bridge". In addition, most words are not accented at all. According to my language consultant, stress does not differentiate different meanings, but rather just makes words and sentences sound subjectively "more natural". To test this claim, I asked Ena to pronounce *shinbunshi* "newspaper paper" while emphasizing a different syllable in the sentence "I did origami with the newspaper".



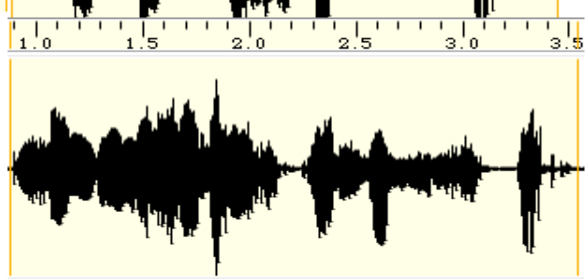


Figure 15. Three different waveforms of "I did origami with newspaper". In sample A (top left), Ena spoke the sentence naturally. In sample B (top right), Ena emphasized *shin* in *shinbunshi*. In smaple C (right), Ena emphasized the final *shi*.

In Figure 15, her sample A, the natural recording, lasted 1.913 seconds, her sample B lasted 2.205 seconds, and her sample C lasted 2.696 seconds. Since B and C were unnatural samples, she tried a combination of syllable lengthening and pitch change, which caused a significant delay in uttering the sentence.

4.2 Phrase Level: How Division of Phrases Relate to the Intonation Contour

Once the accent patterns are determined in the word level, they are basically concatenated together to form the pitch contour of a sentence. Figure 16 shows the example *Tekitōna||bunshō o ||sagashite* "I am looking for a good sentence"[10], where Ena was instructed to divide the sentence into parts she considered phrases. Here, the English ToBI system works well to describe the sample. We get L*L-H*!H-H*H-L%. We can also accurate describe what happens within each phrase: *Teki'tōna* has very minimal stress on the long vowel following the voiceless [i̥] in *ki*. On the pitch contour plot, notice how the end of first blue line is lower than the beginning of the second blue line, the gap being the voiceless vowel. Meanwhile, *bunshō o* has the pattern '--- (the ' denoting stress and the – denoting the other morae), and *sagashite* –'--. Both of these can be predicted from the plot, although the initial *sa* has higher pitch from the transition. Note how *bunshō o* starts with stress and has an exponential-shaped drop for the rest of the phrase. *Sagashite* also drops quickly following its stress.

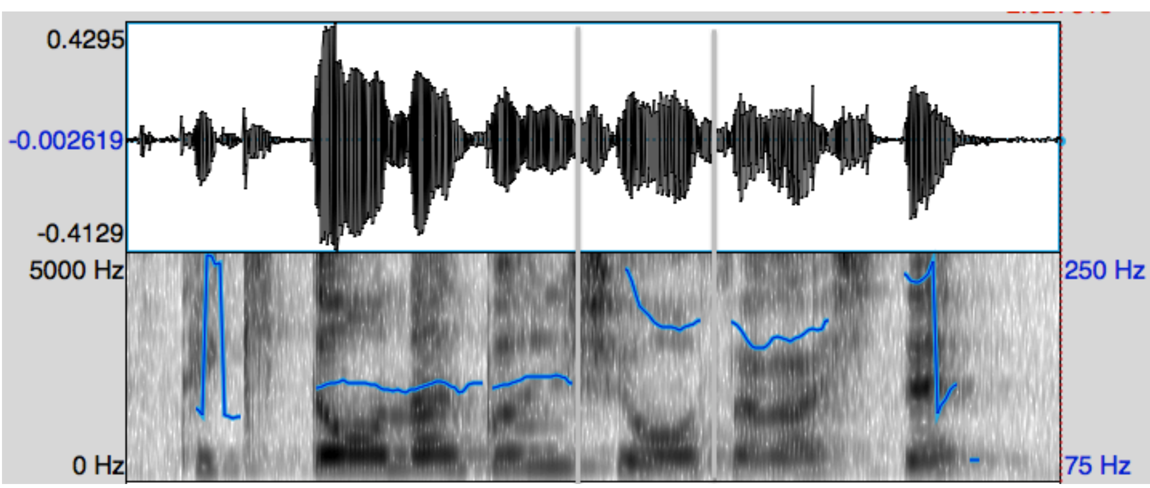


Figure 16. The waveform and spectrogram of ***Tekitōna||bunshō o ||sagashite*** "I am looking for a good sentence". The thick lines in the spectrogram are the pitch contour of the word. The vertical gray stripes are the phrase divisions that Ena made for this sentence.

Now we can begin looking for these patterns in other sentences as well. Figure 17 has the sentence *Kansha-'sai no || shūmatsu wa, || byōin de || otōsan ni || ai ni itta* "On Thanksgiving Day, I visited my grandpa in the hospital"[11]. (Accent patterns: ---'--// '--- // --'-//-'----//-'----). In this example, *shūmatsu wa* has a stress on the first mora, and it shares a similar shape to that of *bunshō o* as we saw earlier. *Kansha-'sai no* from this example resembles *sagashite* from the previous example, and both of them have the stress in the middle of the phrase. It is certainly possible that the accentual pattern can determine that shape of the contour within the phrase. Finally, Japanese speakers have a tendency to rise up and downstep repeatedly in a sentence. In this example, we see a peak in the middle of *Kansha-'sai* and then again in *byōin de* and finally in *ai ni itta*. It would be naïve to conclude that the peaking alternates in the phrases without more sentences as proof, but a correlation may exist just like the stress-within-phrase patterns we just discussed.

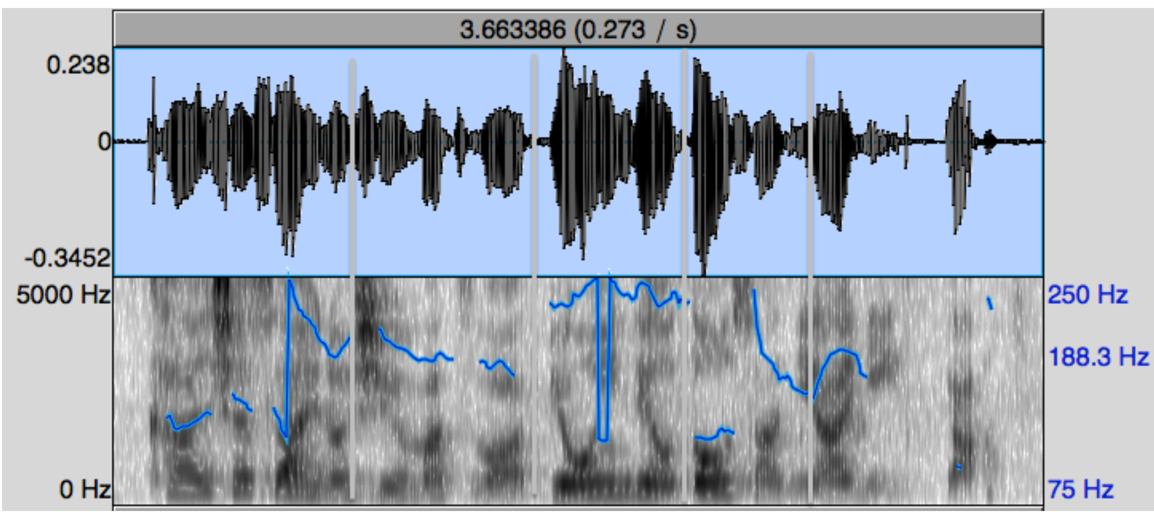


Figure 17. The waveform and spectrogram of ***Kansha-'sai no || shūmatsu wa, || byōin de || otōsan ni || ai ni itta* "On Thanksgiving Day, I visited my grandpa in the hospital"**. The same notation that was used in Figure 16 is also employed here.

4.3 Sentence Level: How Intonation Affects Questions

The last area to explore for intonation is question formation in Japanese[12]. Unlike English, where the intonation of the sentence can change dramatically when converted to a question, Japanese exhibits very little change except for the last phase. Consider the example "Is the sky is blue?" (colloquial) in Figure 18: the question and answer share a similar contour, including most of the peaks, except the very end where the question contour rises and the answer contour falls. The formal question, which includes the –ka particle, exhibits the same behavior,

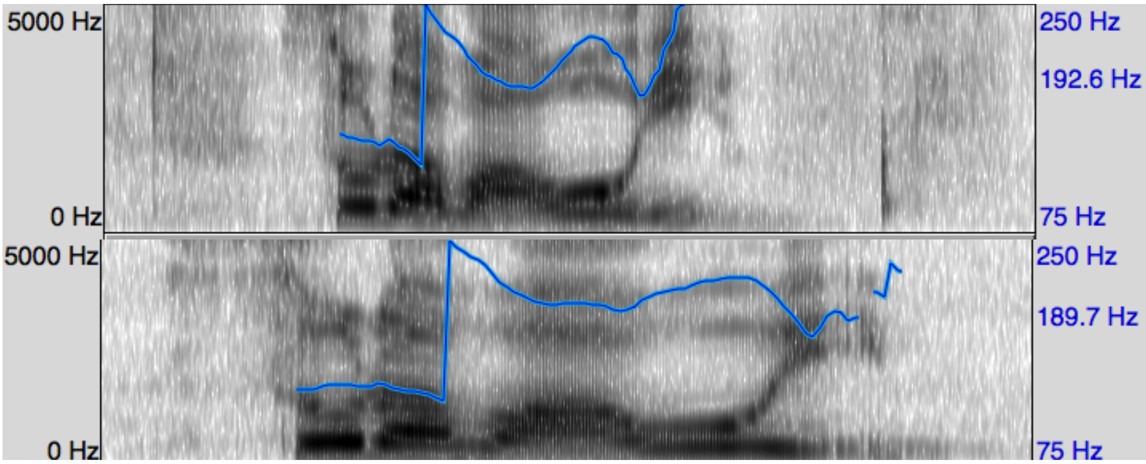


Figure 18. The spectrogram of the question **"Is the sky blue?" (colloquial)** at the top and the response **"The sky is blue"** at the bottom.

although the measurement is skewed due to the fricative and voiceless vowel that comes at the end of the response. Note the sharp rise in the last blue line of the spectrogram, which represents the question being asked, compared against the low starting spot of the answer contour's blue line, which signals the falling intonation of an answer.

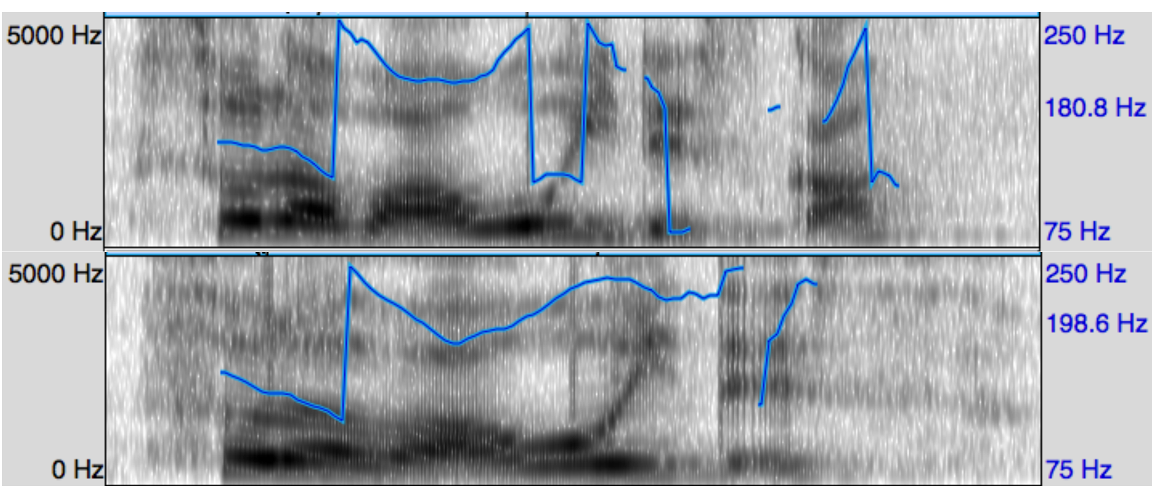


Figure 19. The spectrogram of the question **"Is the sky blue?" (formal)** at the top and the response **"The sky is blue"** at the bottom.

Finally, just as there are one-word questions in English dictated by the intonation – Name? Age? – Japanese has similar expressions. However, upon closer look at the spectrogram, it is clear that one does not have to start low and continue very high in order to ask a question in Japanese. On the contrary, the preceding morae can have high or rising intonation as well, as long as the very last part of the question exhibits some rising intonation. This last-part pitch rise is a recurring pattern in Japanese.

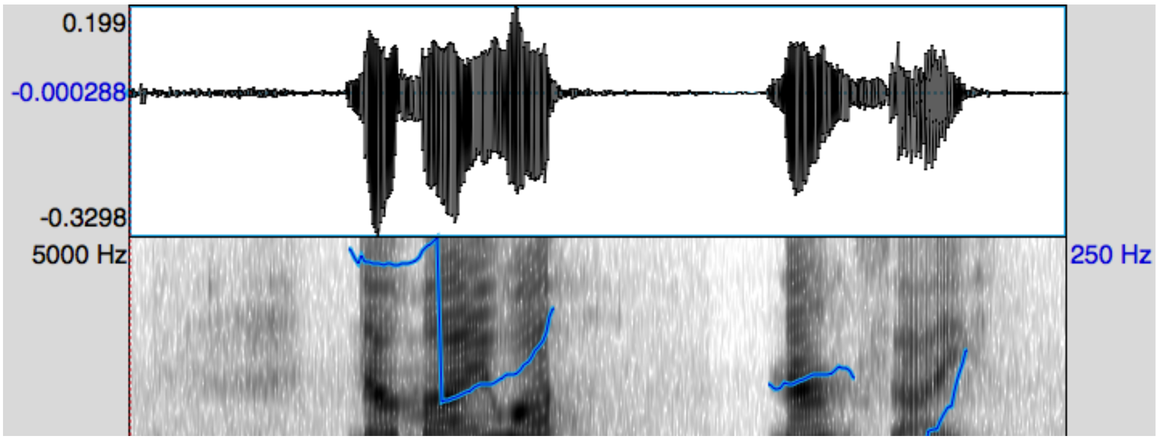


Figure 20. The spectrogram of the question **"Name?"** to the left and "Age?" to the right.

An interesting concluding note is that if one states the polite question (with the –ka particle) without rising intonation, then the question may come off as "direct, ironic, or forceful", according to my language assistant, although it is perfectly acceptable to ask an informal question without any rising intonation.

5. INTRODUCTORY PHONOLOGY

In this section, I will present just a few of the phonological rules I have found while doing research for the other sections with an accompanying word set.

5.1 Place Assimilation of Nasal Coda /n/

Word	Meaning	IPA	Word	Meaning	IPA
hondo	mainland	ˈhõndo	ringo	apple	ɹʲiŋgo
nihongo	Japanese language	ɲiˈhõŋˌgo	nihonhoken	Japanese insurance	ɲihõŋˈhokẽn
jinsei	life	ˈɟ͡ʑĩseː	ane	older sister (coll.)	ane
gan	cancer	gãŋ	nori	seaweed	noɹi
nihonmokunidesu		Japan and Canada		nihõmmokɯnides	
kanli		to manage		kʰãn̠ɹi	

Rule: /n/-> [α place] / __[α place]

The /n/ sound basically assimilates to the place of the following consonant, and possible allophones are [m] in "Japan and Canada", [n] in "life", and [ŋ] in "Japanese insurance". In *kanli,* the [n] is farther back in the alveolar ridge than the [n] in "ane", but not by too much. As noted in consonants section, Ena does not weaken her /g/ ->[ŋ] unlike other Japanese speakers of the Tokyo dialect.

5.2 Palatalization

Word	Meaning	IPA	Word	Meaning	IPA
mi	fruit	mʲi	me	eye	me
hi	son	çi	ha	teeth	ha
ryu	dragon	ɾʲɯ	ruri	lapis	ɹɯɹi
onīsan	brother	õɲiːsãŋ	onēsan	older sister	õneːsãn

Rule: [+consonant]-> [+palatalized] / __ {[i], [j]}

Both [i] and [j] help to palatalize the consonant by making a constriction near the palate, which changes the quality of the sound. /s/ can also become [ɕ] next to an [i]: the Romanization reflects this phonological change and assigns *shi* to /si/. Likewise, /t/ becomes the affricate [cɕ] next to [i], and the difference is reflected in the Romanization. Otherwise, the only palatal consonants are the ones specifically written by the orthography using the letter や. "brother" and "older sister" provide a near-minimal pair, but the change in the [ɲ] does show that the palatalized consonants are in complementary distribution with their regular counterparts.

5.3 Devoicing of [i], [ɯ]

Word	Meaning	IPA	Word	Meaning	IPA
matsu	pine	ˈmatsɯ̥	kikori	woodchopper	kʲi̥ˈkoɾi
futon	futon/bed	ɸɯ̥ˈtõŋ	kōchō	principal	kʰoːˈc͡ɕoː
shinbun	newspaper	ˈɕimbɯŋ	otokō	man	otokoː
moshi	if	ˈmoɕi̥	moshimoshi	hello (colloq.)	'moɕi̥moɕ
hihu	skin	çiɸɯ	tanpopo	dandelion	ˈtʰãmpopo

Rule: [+high (vowel)] -> [-voice] / [-voice –obstruent -stress]__{[-voice], #}

This rule was trickier to write because Ena did not follow as consistent a pattern as the other rules. Nevertheless, this is a common rule found in many Japanese speakers (Akematsu 1997). The main indicator of this rule is the voiceless consonant that precedes the /i/ or /ɯ/, but "dandelion" necessitates the additional obstruent condition. Double vowels do not count as a valid voiceless environment, and stress does play a part in voicing, because it can be difficult to make a higher voiceless sound, given the turbulent flow of those sounds. *hihu*, for instance, does not have devoicing because both morae are stressed.

5.4 Nasalization of Vowel with Nasal Coda

Word	Meaning	IPA	Word	Meaning	IPA
shinbun	newspaper	ɕimbuɯŋ	hon	book	hõŋ
fuusen	balloon	ɸɯːsẽ̞ŋ	hen	transformation	hẽ̞ŋ
mikan	tangerine	mʲikãŋ			
kondo	next time	kõndo			

Rule: [+vowel] -> [+nasal] / __[+nasal]

This rule is commonly encountered because the /n/ is the only nasal coda that exists. Notice also that another general rule is /n/->[ŋ] / __#. Although Akamatsu (1997) explains that the uvular [ɴ] can be found at the end of an utterance, there was no instance where Ena produced an [ɴ] at the end. She says that the tongue is in "resting position" at the end, which implies that there is no effort in moving the tongue all the way back for the uvular sound.

6. CONCLUSION

Even though this research paper covered a lot of ground, from the basic phonetic inventory to the exploration of intonation, there are still quite a few topics that one can explore. For example, the phonology section can be expanded to include the consideration of fast, conversational speech instead of single words. More samples can be taken to get a more accurate measurement of voice onset time for Japanese, and with the possibility of being able to stress any syllable in Japanese under certain constraints, we can explore word accent patterns much more and see how differences in those patterns affect the interaction of the phrases in the pitch contour plots. Further research may include automatic speech recognition (ASR), natural language processing, dialect comparison, and environmental effects on language. (Ena did not grow up in Japan but rather in Michigan with English.)

7. ACKNOWLEDGEMENTS

First and foremost, I would like to thank my language consultant Ena Hariyoshi for volunteering an hour out of her busy EECS schedule every week to help me on this project. 有り難う御座います。 To Florian Lionnet, for his very helpful advice, effective teaching and patience, both in office hours and section, throughout the whole semester. And to Prof. Johnson, without whom this course, and this project, would not be possible.

REFERNCES

Akematsu, Tsutomu. 1997. Japanese Phonetics: Theory and Practice. LINCOM Studies in Asian Linguistics. München: LINCOM EUROPA.

Beckman, Mary E. and Pierrehumbert, Janet B. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook*, Vol. 3, 255-309. Cambridge: Cambridge University Press.

Hasegawa, Yoko and Hata, Kazue. 1992. Fundamental Frequency As an Acoustic Cue to Accent Perception. *Language and Speech* 35, 87-98.

Hasegawa, Yoko. 1995. Against Marking Accent Locations in Japanese Textbooks. *Japanese-Language Education Around the Globe*, vol. 5, 95-103.

Hiragana. 11 December 2013. Wikipedia, the Free Encyclopedia.

Johnston, Eric. 2007. Dialect-rife Japan can be tongue-twisting. The Japan Times. < http://www.japantimes.co.jp/news/2007/11/13/news/dialect-rife-japan-can-be-tongue-twisting/#.Uqo8PWRDuLk>

Ladefoged, Peter and Johnson, Keith. 2011. A Course in Phonetics (Sixth Edition). Boston: Wadsworth, Cengage Learning.

Lewis, M. Paul; Simons, Gary F.; and Fennig, Charles D. (eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International.

Mannell, Robert. 2008. Acoustics of Speech: Speech Spectra and Spectrograms 6.e: Approximants. Sydney, Australia: Macquarie University.

Nationalencyklopedin. 2010. "Världens 100 största sprak 2010" (The World's 100 Largest Languages in 2010).

Otaka, Hiromi. 2009. Phonetics and Phonology of Moras, Feet, and Geminate Consonants in Japanese. Lanham, Maryland: University Press of America, Inc.

Ryan, Camille. 2013. Language Use in the United States: 2011. American Community Survey Reports. U.S. Census Bureau. Accessed at http://www.census.gov/prod/2013pubs/acs-22.pdf

Shibles, Warren A. 1994. Chinese Romanization Systems: IPA Transliteration. Sino-Platonic Papers No. 52. Philadelphia: University of Pennsylvania Department of Asian and Middle Eastern Studies. Accessed at http://www.sino-platonic.org/complete/spp052_chinese_ipa.pdf

NOTES

[1] From Table 1 of the report *Language Use in the United States: 2011* (Ryan 2013).
[2] At the beginning of my research, my language consultant did not consider it natural to investigate the individual sounds of a syllable. Nevertheless, those individual sounds should be considered for purposes of classifying the different types of onsets and forming phonological rules.
[3] See Akematsu (1997 p. 160) or Otaka (2009 p. 2) for a deeper discussion about the validity of the "mora hypothesis". Ladefoged and Johnson (2011 p. 251) summarize the concept nicely.
[4] I chose to use [ɯ] instead of [ʉ] despite the results from Figure 2, the vowel plot from the morphemes, because Japanese does not have a contrast between [ʉ] and [ɯ], and in addition, the vowel is unrounded. Refer to Akematsu (1997 p. 284) for a similar discussion. The plot measurement may be skewed by my consultant's particularly forward pronunciation in "su", with an F2 of 1883.3 Hz versus "ku", with an F2 of 1536.6 Hz and her held-out vowel sound, which measures at 1316.97 Hz. In her pronunciation of "gu" (ingredients), her F2 is 1631.87 Hz. Although Wikipedia uses the more accurate [ɯ$^{\beta}$], where the $^{\beta}$ represents the compression of the lips, $^{\beta}$ itself is not defined in the IPA, so I will not use that symbol here.
[5] The other example is *Tōkyō-do*, where the ō Romanization obscures the fact that the Japanese spell the sound as *ou*. However, this example is just orthography-related, as my language consultant considers it phonemically as /oː/.
[6] The transcription of the Chinese sound *x* (Pinyin) is from Shibles (1994). Example words are *hsü* (Wade-Giles)/*xu* (Pinyin). Shibles' list is alphabetized by Wade-Giles. I recognize this sound as a native speaker of Mandarin Chinese.
[7] According to my language consultant, and supported by the Wikipedia page on Hiragana.
[8] My interpretation of Beckman and Pierrehumbert's conclusion (1986). Hasegawa (1995) interprets the paper similarly in her Section 4.
[9] Akematsu (1997) dubbed this pitch phenomenon a "plateau", and considers the stress assignments as "accent patterns".
[10]適当な|文章を|探して。
[11]感謝祭の|週末は、|病院で|お父さんに|会いに行った。
[12] I tried asking Ena about other applications of intonation, like how telephone numbers and mathematical operations are read, but she said that there was very little difference from a straight read-through except for the pauses. The spectrograms show that her reading is very similar to English in terms of intonation, so no further study will be done on those topics.