

FRANCESCA PINTO

High vowels elision in Japanese. A perceptive approach

The present study will focus on the perception by the native speakers of elided high vowels in Japanese and of deleted sounds in spontaneous speech. Study of mother tongues perception of the said phenomenon may serve as a capstone toward the acceptance of elided vowels in Japanese. It is a common belief among scholars that vowel elision may interfere with words comprehension, for the missing vowel would create a high rate of ambiguity, especially when dealing with minimal pairs. For this reason, the present study aims to demonstrate that words are easily understood despite vowel elision, through a perceptive test based on the recognition of 11 minimal pairs. Results revealed that vowel elision does not cause much trouble in the identification of the target words.

1. *Introduction*

Japanese high vowel devoicing is a well-known and studied phenomenon, examined by many native and non-native scholars¹ since the early stages of literature on Japanese linguistics.

High vowel devoicing is a phenomenon for which vowels are pronounced without the vibration of the vocal folds and where oral organ configuration is more similar to that of whispering: they are hence voiceless. A voiceless vowel can be easily distinguished from a fully voiced vowel for a series of acoustic peculiarities: it does not display any trace of the voice bar at the bottom of the spectrogram, it shows a sensitive increase of both F1 and F2, formants look less neat and sharp, pitch is absent, signal is aperiodic, there should be relatively little or no turbulent airflow during articulation², and overall the phone length is shorter. Vowel devoicing is more likely to occur in specific phonotactic contexts, specifically between two voiceless consonants or between a voiceless consonant and a pause. Factors like speech rate may also have a strong influence.

Under the same circumstances, high vowel elision may occur as well. Acoustically, a vowel is said to be elided when there is the complete absence of the phone with all its characteristics and therefore cannot be retraced. Depending on a higher speech rate, on the quality of the consonantal environment, and on the lexicon recurrence, this latter phenomenon may have an even higher degree of occurrence as compared to vowel devoicing. In fact, it is quite evident that some specific contexts require vowel elision more than fully voiced or devoiced vowels.

¹ See, for example, Akamatsu (1997), Maekawa (1993), Shibatani (1990), Vance (2008), Yuen (2000).

² This is needed to distinguish devoicing from frication, Yuen (2000: 532).

The reason for the alternation among vowel devoicing or elision is partly due to the ease of pronouncing certain consonantal clusters instead of others. A cluster (resulting from the elision of an interconsonantal vowel) composed, e.g., of a palatal fricative plus an alveolar occlusive is easier to pronounce than one composed of two occlusives. Additionally, word recurrence plays an important role in determining vowel elision regularity. Words like *desu* 'to be' or *shita* '(verbal morpheme for past tense)' are very frequent in Japanese and their realization with high vowel elision has gradually become a common habit of pronunciation which conflicts with the written interface of those words.

Notwithstanding neat phonetic evidences that prove vowel elision, traditionally only vowel devoicing is accepted by scholars in this field, especially by native ones³. It is possible to assume that the reason for that is mainly due to the influence exerted by Japanese writing system. Apart from Chinese ideograms and Latin characters, Japanese language is represented by *kana*, a syllabic writing system where each symbol stands for a mora composed of a vowel or a consonant plus a vowel, i.e. an open syllable. The only exceptions to this are represented by the symbol <n>, the only consonant that can occur alone and form a mora, thus closing the syllable, and by the symbol <tsu> conventionally used in a smaller size to mark intense consonants, which makes the preceding syllable closed as well.

As a consequence, it is not rare that native speakers experience difficulty in discriminating acoustic differences between vowel devoicing and elision. In other words, Japanese *kana* has influenced the perception of Japanese phonology and phonotactics so much that even scholars are led to think that only open syllables are possible – with the said exceptions – and hardly acknowledge vowel elision and consonantal clusters that derive from this process. Therefore, since speakers' intuition is not reliable in this case, a direct observation of acoustic data is needed.

The aim of the present study is to understand to what degree Japanese mother tongues' intuition, notoriously influenced by the writing system, is in contrast with the acoustic and perceptual reality. Investigating perception represents a key point to the full understanding of this research topic and of the possible drift of the language. This research may also constitute an original contribution to the research in this field, since perceptive tests are still not very popular in this branch of research.

Unfortunately experimental analyses on speaker perception are still not very popular in this field. On the other hand, investigating the native speakers' way of interpreting speech is an efficient method to confirm results from acoustic analysis, especially when interacting with phenomena that may be suspected of being in the middle of an ongoing linguistic change. What will be exposed here is partially examined in another study, dealing with the analysis of acoustic and

³ See, for example, Akamatsu (1997: 50).

perceptive aspects of vowel reduction in Japanese, which has been conducted by Yuen⁴.

Yuen starts from the acoustic analysis of six minimal pairs pronounced by eight native speakers – the large majority of the analysed vowels were elided – and successively used the same tokens as the basis for a perceptive test, which he conducted on forty native speakers. The experiment conveyed by Yuen hence originates from a different analysis basis, not only due to the number of tokens and subjects taken into account, but also for the fact that he considers also devoiced – in addition to elided – realizations of high vowels. On the other hand, in the present study only elided vowel's role will be analysed, since it is believed that devoiced vowels affect word recognition, despite being reduced. Notwithstanding this difference, it is particularly relevant that the author got experimental results that are similar to those illustrated in the present paper: even though most of the devoiceable vowels are deleted, the listeners did not seem to have problems recovering the underlying vowel identity.

Apart from the recognition of word including elided vowels, the present paper aims to demonstrate that vowel elision can be considered as an innovative feature in Japanese, also on the base of relevant acoustic analyses on diachronic and diatopic variants⁵. In fact, it is possible to postulate that an elided vowel may have consequences on Japanese's phonological system, mainly dealing with its palatalization correlation. As results will show, the correct understanding of the proposed stimuli occurs through coarticulation phenomena on the preceding consonant. It is well known that coarticulation is crucial in phone recognition and in the present experiment it plays a key role in determining the correct identification of the proposed stimuli, even though it occurs at different degrees according to the nature of the consonant. This process constitutes a matter of importance, because phonetic traces deriving from elided vowels allow the correct transmission of the message informative content with no comprehension inconvenience. The said phenomena, being a constant feature of consonants followed by front vowel, would enrich the already existent correlation, opposing consonant followed by non-front vowels to the palatalized ones.

1.1 Expected results

Expected results will show that the listener will be able to correctly identify most of the proposed words, notwithstanding vowel elision. The reason for this assumption relies on the fact that phonetic traces still remain in the coarticulation of the preceding consonant. These traces will constitute the necessary hint for the right comprehension of the proposed stimuli.

For example, if we take the pair *kisai* and *kusai*, we can expect that the speaker will decode the proposed input due to the different pronunciation manner of

⁴ Yuen (2000).

⁵ See, for example, Byun (2011), Frellesvig (2010), Maekawa (1993), Shibatani (1990).

/k/, [k̥] in the first case and [k] in the second, coherently with the palatalization correlation already present in this language.

2. *Methodology and collected data*

2.1 Subjects

Fifty Japanese native speakers with no apparent auditory problems participated as volunteers. Other than a privacy and a safety agreement form, subjects were asked to fill in an anonymous form dealing with information about their personal and linguistic history such as: sex, age, birthplace, living place, number of years spent in the living place, mother's birthplace, father's birthplace, and highest education level obtained.

The information required was needed because it is not rare that subject's pronunciation may vary greatly according to sociolinguistic factors. In particular, diatopic related information was needed in order to understand the degree of influence of the dialect on missing vowels discrimination. It is a well-known fact⁶ that the phone reduction phenomena may vary according to the spoken language geographic variety. For example, Japanese speakers from the Kansai area⁷ are more used to the rounded high back vowel, which is more audible and nearer than its unrounded counterpart, typical of the Standard Japanese.

On the other hand, information related to the education level was needed in order to evaluate the degree of awareness of the first language speakers towards their native language. In fact, higher educated speakers are considered to have a more accurate knowledge of language mechanisms.

Table 1 - *Speakers dialectal background related data*

	<i>Birthplace</i>	<i>Mother's birthplace</i>	<i>Father's birthplace</i>
Hokkaidō	0	0	0
Tōhoku	2	2	2
Kantō	35	24	24
Chūbu	9	8	8
Kansai	1	1	0
Chūgoku	2	1	1
Shikoku	0	0	0
Kyūshū/Okinawa	0	0	0
Other	1	0	0

⁶ Maekawa (1993).

⁷ Kansai area includes Kyōto, Ōsaka and Kobe dialects, among others.

⁸ In the following table data related to mother and father birthplace were reported only when they coincide with those of the speakers' birthplace. In this way it will be easier to verify the dialectal influence of the parents' birthplace on the subjects.

Table 2 - *Personal data related to the 50 speakers*

Sex		Average age	Years spent in Tōkyō	Education level ⁹			
M	F			H	B	M	P
46	4	24,64	16,33	13	28	6	3

2.2 Stimuli and task

The test is based on a series of 11 minimal pairs (i.e. 22 lexical items). Selected words are extracted from an interview expressly created for this research purpose. A native speaker with no apparent pronunciation problem or strong accent was asked to pronounce each word with the elided pronunciation of the vowels involved. Words were then carefully checked one by one with Praat to make sure that each target vowel was elided. Words were chosen according to the vowel quality and to the consonantal environment. Furthermore, particular attention was paid in order to have a similar speech rate and not to combine rare and common words so that the listeners would not be influenced by the lexicon occurrence rate. It was verified that all the subjects were familiar with the chosen words.

Pairs including meaning differences caused by /i/ and /u/ in the same position within the word were selected so that vowels occurred both in low pitched and in unaccented morae. Due to a palatalization process occurred in the 16th century¹⁰, many voiceless consonants shifted place or manner of articulation or both in front of high vowels, producing what are now complementary distributed allophones, e.g.:

/t/+/i/ → [tɛ], vs. /t/+/u/ → [t̪] ([t] otherwise);

/h/+/i/ → [ç], vs. /h/+/u/ → [ɸ] ([h] otherwise)¹¹.

This is the reason why only [k] and [ç] are used here, for they are the only voiceless consonants that can contrastively occur before both high vowels (note that, while [ɛi] is still a result of the palatalization of /s/+/i/, it is nevertheless contrasted with [ɛu] coming from the palatalization of the cluster /si/ followed by /u/). Considering this, and in order to match minimal pairs in the most convenient way, all the other voiceless consonants have been excluded. As for the following consonant an attempt was made to vary the possible phonotactic contexts.

Selected words are the following:

⁹ H stands for High school; B stands for Bachelor degree; M stands for Master degree; P stands for PhD.

¹⁰ Frellesvig (2010: 322).

¹¹ Japanese is among the few languages that allow palatalization phenomena in front of a back vowel. See Bath (1978: 53).

Table 3 - *Selected pairs*

Word	Transcription	Meaning	Pair	Transcription	Meaning
<i>Shishoku</i>	/cieokuu/	Sampling food	<i>Shushoku</i>	/euæokuu/	Staple food
<i>Kishin</i>	/k'iein/	Donation	<i>Kushin</i>	/kuæein/	Effort
<i>Ishitsu</i>	/ieitsu/	Heterogeneous	<i>Ishutsu</i>	/ieutsu/	Export
<i>Kakite</i>	/kak'ite/	Writer	<i>Kakute</i>	/kakute/	Thus
<i>Yukisaki</i>	/juk'isaki/	Destination	<i>Yukusaki</i>	/jukusaki/	Future
<i>Kikai</i>	/k'ikaj/	Machine	<i>Kukai</i>	/kuukaj/	Recital
<i>Kisai</i>	/k'isaj/	Mention	<i>Kusai</i>	/kuusaj/	Smelly
<i>Kichō</i>	/k'iteo:/	Precious	<i>Kuchō</i>	/kuuteo:/	Tone
<i>Shikan</i>	/eikan/	Official	<i>Shukan</i>	/euukan/	Subjective
<i>Shitai</i>	/eitaj/	Corpse	<i>Shutai</i>	/euutaj/	Subject
<i>Shichō</i>	/eiteo:/	Mayor	<i>Shuchō</i>	/euuteo:/	Opinion

For the aim of this test subjects were asked to listen to each of the 22 words and to choose between two provided alternatives forming the minimal pair, namely the word containing /i/ and the one containing /u/. Overall, subjects needed one to three hearings per each word. Total time needed to complete the test was less than three minutes.

2.3 Analysis of data

Words pronounced by the native speaker within a sentence have been cut and then proposed to listeners without the help of a phonetic or a semantic context. This process was needed so that the listener would have not been influenced by any other element that might have interfered with words identification.

For word segmentation the Praat software, version 5.3.49, has been used. The test has been developed using the Max Runtime software, version 6.1.7. Metadata was installed into the said software and selected words were reproduced randomly. The said software was programmed in order to obtain a yes/no output for each audio segment, corresponding to the correct or incorrect guess of the word including the elided vowel. Test answers were automatically collected in a database and statistically analysed using the correction for guessing formula¹², a method that clarifies the link between the score on the test and the subject's real understanding.

The test was run with the support of Sony stereo headphones MDR-XD100 and was conducted in a soundproof room.

¹² Chiorri (2011: §4.7)

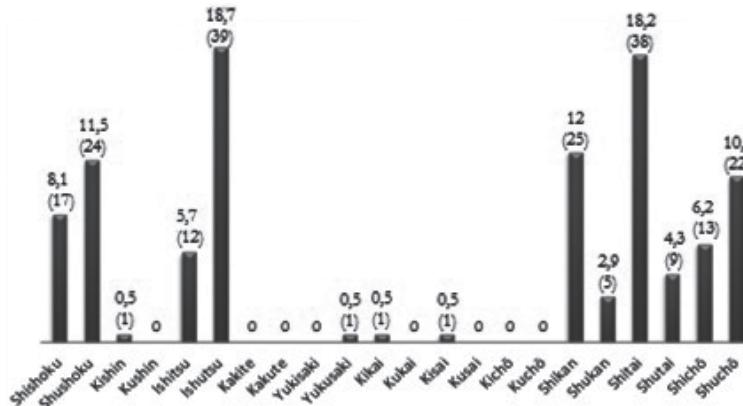
3. Test results

Here follow tables and graphs that illustrate results from the perception test.

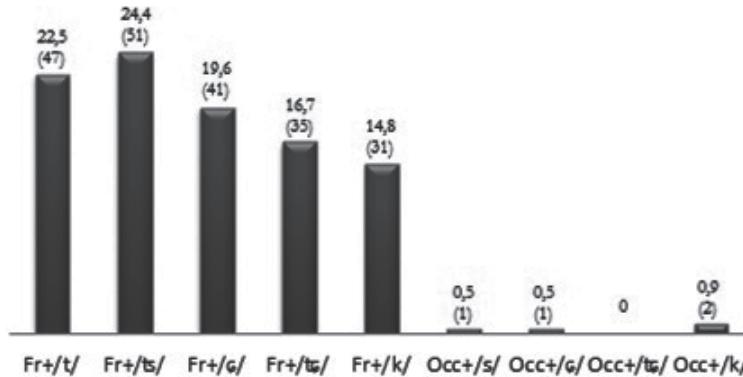
Table 4 - *General results*

<i>Tot. amount of proposed stimuli</i>	<i>Tot. amount of identified words</i>	<i>Tot. amount of mistaken words</i>
1100	891	209
100%	81%	19%

Graph 1 - *Word related error rate plus error total amount in brackets*



Graph 2 - *Phonotactic context related error rate plus error total amount in brackets*



Results from the test based on Japanese minimal pairs seem to confirm the above mentioned hypothesis according to which native speakers can correctly discriminate words involving high vowel elision.

The total amount of proposed stimuli was 1100 and the total number of errors committed was 209. Using the correction for guessing formula and considering that

test responses might have been the result of a random guess, it is possible to state that the total amount of correctly guessed words is 62%.

As graphs 1 and 2 show, the number of errors differs highly depending on the phonetic context, but the question is which part of the context is more influential. It seems that the timbre of the vowel involved does not have a significant influence on the error rate for the number of errors related to /i/ is 108 and 101 related to /u/. Nor does the following consonant have a strong influence on the error production. Not surprisingly, it is the preceding consonant that exerts a most considerable impact on the occurrence, which is due to some very clear phonetic reasons. This is quite evident when comparing error rate related to words where an elided vowel is preceded by a fricative to those preceded by an occlusive, for coarticulation reasons that will be explained later; while in the first case the number of errors is 205, in the latter case it is only four.

Another point of interest is represented by the eight out of twenty-two words that were never mistaken, i.e. *kushin*, *kakite*, *kakute*, *yukisaki*, *kukai*, *kusai*, *kichō*, *kuchō*, which are the ones that involve [k], regardless of the following vowel or consonant. Even the mistaken words where [k] is the preceding consonant were misunderstood only once per each word. On the other hand, ambiguity caused by [č] is much more relevant, even though the number of errors differs significantly depending on the word. The lowest number of errors, 6, is related to the word *shukan* 'manager' while the highest, 39, goes with the word *ishutsu* 'export': using the correction for guessing formula and assuming that the maximum error rate here is 50%, for the subjects only had two alternatives, the error rate of the word *shukan* is 76%, while the error rate of *ishutsu* is 56%. There are no single words mistaken by all participants.

It seems quite evident at this stage that coarticulation caused by elided vowels has a different effect depending on the quality of the preceding consonant. Stops seem to better preserve coarticulation – a clear example of this is represented by aspiration that usually follows the phone burst when combined with /i/ – making words identification much easier. On the contrary, coarticulation on fricatives mostly disperses and only a few traces remain: this is the main reason why a higher degree of ambiguity occurs in coincidence with these phones.

For what concerns subjects' related data, in this test the average number of errors is about four per speaker, with a minimum of two errors and a maximum of seven. The majority of the test subjects were from Tōkyō and surroundings, so it is possible to say that this sample may represent a good mirror of the perception of this phenomenon in the standard variety of Japanese (which correspond to the dialect of the capital). Dialectal differences do not seem to have a great influence on perception as answers given by speakers from other regions are totally comparable to those given by Tōkyō speakers.

4. *Discussion*

Results from this test are quite interesting for two main reasons: they constitute new support material for the hypothesis according to which Japanese phonology may eventually include elided vowels and consonantal clusters that would derive from it and they can contribute to the advancement of the state of the art of missing phones recognition in speech.

These results show that elided vowels are well tolerated perceptually. In general, we can state that notwithstanding differences in the degree of identification of words including vowel elision, missing phones do not consistently obstacle word recognition and do not cause any relevant trouble to the message transmission. We should not forget that in spontaneous speech and in everyday language the number, type of errors, and ambiguities would be even lower, due to the role of semantic context, of sentence intonation and of all the other mechanisms involved in oral communication, whereas in the present study words were proposed in isolation. Moreover, in spontaneous speech phones elision is quite frequent, especially when the speech is inaccurate and at a high rate. For these reasons, speakers already have the habit to interpret fragmented speech.

It is possible to state that when speaking, people are more interested in the mere pragmatic aspects of the language and the message transmission is a priority over the speech hyperarticulation. It frequently happens that in a linguistic act, especially when inaccurate, there is an apparent high number of imperfections and gaps: it is not rare that speakers disregard the speech form and orthoepy and prefer to focus more on other aspects of the communication. In this way elisions or abbreviations often take place, being them of whole words or small fragments. When there is a good mutual comprehension, speakers will also benefit from an articulatory and mental saving, without drawing upon an orthoepic and accurate speech. In other words, during a speech act we do not hyperarticulate sounds, but we rely upon conversational redundancy. This allows us to economize on articulatory effort when we can count on listener's cooperation.

The dialectic balance among interlocutors finds its basis on the principle of pragmatic cooperation principle¹³. According to this principle, complicity and participation are established among people involved in a conversation, so that orthoepy no longer represents the most important parameter. Speakers, in fact, capitalise on their lexical competence and activate a meaning recognition mechanism – when the discourse is pertinent – in order to trigger phonetic perception. Phone elision is certainly able to generate ambiguities, especially as related to minimal pairs, in languages with a high rate of homonymy. The context then, plays a fundamental role in determining the semantic field of a conversation in case of ambiguities. Hence, the listener will not be led to identify the missing phone with the wrong vowel, but will understand without many efforts the pronounced words, even though an inaccurate pronunciation or a high speech rate might have provoked vowel elision.

¹³ Grice (1975: 45).

4.1 The role of coarticulation in speech recognition

As mentioned above, one of the most important perplexities raised by scholars in this field deals with speech recognition, or to put it into other words, the impossibility to correctly identify words including vowel elision¹⁴. The test presented above showed that stimuli recognition, especially when detached from the context, occurs mostly due to vowel coarticulation – even when not pronounced – on the preceding consonants, representing somehow the last trace of its presence.

Results from the perceptive test on mother tongue speaker's perception showed that in the vast majority of cases subjects are able to discern minimal pairs of words where the main obstacle was represented by the elision of the vowel involved. The given explanation was that subjects focused their attention to the acoustic environment of the elided vowel – mostly on the preceding consonant – to clearly label each proposed stimulus. In other words, we can state that the effect of coarticulation on the preceding consonant might have been the main feature for the subjects to give the correct answer. This process takes place because a linguistic chain is not made by single sounds, separated among them, pronounced and perceived in isolation. We can state instead that sounds tie up¹⁵.

Hence, in the missing phone recognition process and, more specifically, when dealing with Japanese high vowels, coarticulation plays a very important role¹⁶. In fact, as stated above and as Kawakami¹⁷ notes, vowel identification, even though elided, occurs at any rate through coarticulation clues exerted by /i, u/: more specifically, it ensues in a more advanced position in the first case and in a more retracted one in the second case. We can expect that the said mechanism takes place also in conjunction with a devoiced vowel, as well as with a fully voiced one. The fact that this may take place also coincidentally with an elided vowel is something that raised experts' doubts and perplexities, whose most consistent objection is that, summing up what Kondo¹⁸ says, if a vowel no longer exists, it shall not cause coarticulation. Even though we can state that a missing sound may no longer provoke an influence on the phonotactic context, not being part of it anymore, we should not forget that vowel elision is the result of a process for which a vowel undergoes many reductions going from a status of full sonority to its complete elimination. Note that during this process the subsequent consonant can retain the coarticulation effects, thus leading to the split of the allophones.

What has been said can give elucidations about some scholars doubts – among them Akamatsu – according to whom the recognition of words including vowel elision might be highly undermined by this phenomenon and the ambiguity degree would be even higher when dealing with minimal pairs. In fact, if we consider the coarticulation impact, word related phonetic image would be certainly differ-

¹⁴ See for example Akamatsu (1997: 50).

¹⁵ Nusbaum, DeGroot (1990: 296).

¹⁶ Labrune (2006: 51).

¹⁷ Kawakami cited by Vance (1987: 54).

¹⁸ Kondo (1997: 14).

ent and this is quite evident if we try and give a narrow phonetic transcription of the words considered here. For example, if we take the pair *kishi* 'shore' /kici/ and *kushi* 'comb' /kuei/, in case of vowel elision we would not have, as many scholars say, a pronunciation like [kei] in both cases, but [k'ei] in the first case and [kei] in the second. To put it into other words, a slightly more advanced pronunciation of the occlusive is the key that would grant speech recognition without many efforts. As Vance states¹⁹, the phoneme /k/ in *kishi* is realized with the allophone expected in front of /i/, while the /k/ in *kushi* with the allophone expected in front of /u/.

It shall be noted that at the present stage of Japanese, apart from the exception of /n/ and of geminates (in case we interpret them as clusters of identical consonants), its phonological system only allows the presence of a vowel preceding or following a consonant. Therefore, if we pronounce the phoneme /tɕ/ it will necessarily follow /i/; in the same way, after the phoneme /ɸ/ there must be /u/²⁰.

As mentioned above, Japanese phonotactics is mainly composed of CV sequences and it goes without saying that the presence of an elided vowel would lead to relevant modifications of the simply panorama of its phonotactic rules. We may postulate, in fact, an increase of possible combinations of consonants and vowels and above all a wider number and types of consonantal clusters. This process is absolutely coherent with the present system and with the naturalness of the linguistic types. Even though this is not the place where to exhaustively discuss this topic, we can briefly postulate the creation on the following syllable structures: (1) VC; (2) CVC; (3) CCV; (4) CyVC; (5) CCVN; (6) CCVQ; (7); CCVR.

4.2 Possible effects of high vowels elision on the palatalization correlation

In the introductory paragraph it was mentioned that we are probably observing a mutation process in act, regarding the acceptance and the inclusion of the phonological consequences that would arise from vowel elision, primarily the increase of the quantity and the complexity of consonantal clusters.

It was said that coarticulation played a fundamental role in stimuli recognition. Coarticulation may sometimes induce cases of allophony, most of the time due to the influence of adjacent sounds in the speech chain: in some cases new forms prevail and become phonologically stable. A further step is when the two allophones split into two phonemes because of the elision of the conditioning context, such as the vowel elision under consideration here.

It is well known that allophones participate in speech recognition in that they anticipate the following phone, for example, as briefly seen above, [ç, tɕ, dʐ] precede [i], as much as [ɸ, tʂ, dʐ] are connected with [u], being the allophones of /h, t, d/. Since reduced vowels cause coarticulation on the preceding consonants and being it a steady phenomenon, it is possible to state that they may originate new allophones. If we take a closer look at the sequence [ki], taken into account for the experiment

¹⁹ Vance (2008: 209).

²⁰ Exceptions deriving from loanwords allowing other combinations are not taken into account here.

presented here, we will see that it contemplates an aspiration after the VOT which is intrinsic to its pronunciation, making this sound tend to palatalization. The said characteristic allows this sound to be clearly distinguished from occurrences where /k/ is followed by other vowels²¹. Moreover, aspirated realization of /k/ in front of /i/ is a constant and regular phenomenon: it is not an irregular or a sporadic circumstance, but it has a peculiar stability. Even from a perceptive point of view, as the test results show, the identification of the said articulation when not followed by a vocalic phone takes place through the acoustical recognition of the frication accompanying the consonant, making values similar to those of a palatal consonant. Because of this recognition process, subjects will overtake the potential conditioning caused by the image of the writing system that a speaker may take as a datum point.

Given the identification degree, it is possible to postulate that in this language the palatalization correlation is enlarging, with the inclusion of new palatalized consonants. In other words, we can hypothesize an increase of the correlation through the phonologization of the allophone [k'], which is already existent in this language. Supposing that high vowels elision might become a more stable reality, it can lead to the consequence of the phonologization of the palatalized allophone of /k/. In support of this hypothesis, it should be reminded that at the present stage of this language the palatalized variant of /k/ already appears in front of /a, u, o/ and in front of /e/ in case of loanwords. This sound, hence, is not only existent in Japanese, but is also part of the speakers' phonetic and articulatory habits.

It is a matter of fact that languages are subjected to constant changes which are most of the time coherent with its system, in other words, they follow their drift. Languages are economical systems where nothing is superfluous: when changes occur they have the function to balance what other changes has made unstable. As Lazzeroni says²², since every system tends to its balance, when a language experiences some irregularities, balance will be restored through other changes. It is clear that the unbalanced situation caused by vowel elision may be successfully rebalanced through the enlargement of palatalization correlation, other than a new way of structuring morae and syllables. When a phonological system already displays that correlation, there are higher chances that it will be enriched through consonantal phonemes not yet integrated in it. The author also illustrates²³ that the presence itself of an opposition leads the system to extend that distinctive feature to more phonemes, through a new modification.

At this point, it is possible to presume that the influence of the elided vowels on the preceding consonant may lead to the creation of new minimal pairs, through new palatalized consonants that would enhance the already prolific palatalization correlation in this language. At the moment these palatalized consonants shall be considered as context determined allophonic variations, since within speech from

²¹ A clear example of this, pertaining to the couple *kishi* – *kushi*, is exposed in the previous paragraph.

²² Lazzeroni (1998: 50).

²³ Lazzeroni (1998: 46).

the same speaker, realizations of the same lexemes occur both with elided or devoiced vowel, according to the diaphasic variance and, consequently, to the speech rate. Hence, we must suppose that, from a paradigmatic point of view, these vowels are phonologically present for the speakers. However, in case vowel elision would predominate in the course of time in a consistent group of speakers, we should hypothesize that a phonologization process of the palatalized allophones of these consonants might have been established. The said process would lead to an enrichment of the already existent palatalization correlation in Japanese phonology. Even though these consonants are not phonologized yet, these allophones constitute a good prerequisite to hypothesize that this phonologization process may eventually take place in Japanese.

Acknowledgements

The author would like to express her gratefulness to Prof. Claudia Ciancaglini, Prof. Andrea Paoloni, Prof. Artemij Keidan, Prof. Shinoda Koichi, Prof. Maekawa Kikuo, Prof. Timothy J. Vance and Prof. Kibe Nobuko for their precious comments and suggestions. Many thanks also to Prof. Ozawa Naoko for her assistance with the Japanese based test and to Dr. Massimiliano Todisco from the Fondazione Ugo Bordoni that helped with the engineering part of the experiment. The author is also grateful to the fifty subjects that voluntarily participated in this research.

References

AKAMATSU, T. (1997). *Japanese Phonetics: Theory and Practice*. Munchen, Newcastle.

BHAT, S.D.N. (1978). A General Study of Palatalization. In GREENBERG, H. (Ed.), *Universals of Human Language*, vol. 2. Stanford, California: Stanford University Press, 47-92.

BYUN, H. (2011). High Vowel Devoicing in the Keihanshin Area of Japan. In *The Journal of the Acoustical Society of Japan*, vol. 15 n. 2, 23-37.

CHIORRI, C. (2011). *Teoria e tecnica psicometrica. Costruire un test psicologico*. McGraw-Hill Companies S.r.l., Publishing Group Italia.

DUPOUX, E. et al. (1999). Epenthetic Vowels in Japanese: a Perceptual Illusion? In *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568-1578.

FRELLESVIG, B. (2010). *A History of the Japanese Language*. Cambridge, University Press.

GRICE, H.P. (1975). Logic and Conversation. In COLE, P., MORGAN, J. (Eds.), *Syntax and Semantics*, vol. 3. New York: Academic Press, 41-58.

KONDO, M. (1997). *Mechanism of Vowel Devoicing in Japanese*, unpublished PhD thesis.

JAKOBSON, R. (1973). *Essais de linguistique générale*. RUWET, N. (Ed.), Paris: Les éditions de minuit.

LABRUNE, L. (2006). *La phonologie du japonais*. Leuven, Paris: Peeters.

LADEFOGED, P., JOHNSON, K. (2010). *A Course in Phonetics*, 6th edition. Cengage Learning.

LADEFOGED, P., BROADBENT, D.E. (1957). Information Conveyed by Vowels. In *The Journal of the Acoustical Society of America*, vol. 29, n. 1, 98-104.

MAEKAWA, K. (1993). Boin no museika. In MIYAJI, S. et al. (Eds.), *Kōza nihongo to nihongo kyōiku, 2 Nihongo no onsei – on'in*, vol. 1. Tōkyō: Meiji Shoin, 135-153.

NUSBAUM, H.C., FRANCIS, A.L. & LUKS, T. (1997). Speech Perception: A Special Mechanism or a Specialized Cognitive Process? In *CLS 33*, vol. 2, 415-434.

SHIBATANI, M. (1990). *The Languages of Japan*. N.Y.: Cambridge University Press.

VANCE, T.J. (1987). *An Introduction to Japanese Phonology*. Albany, N.Y.: State University of New York Press.

VANCE, T.J. (2008). *The Sounds of Japanese*. Cambridge: University Press.

VYGOTSKY, L.S. (1980). *Pensiero e Linguaggio*. MASSUCCO COSTA, A. (Ed.), Firenze: Giunti.

YUEN, C.L.K. (2000). The perception of Japanese Devoiced Vowels. In *CLS 36*, 531-547.