# WPLC

Working Papers of the Linguistics Circle
of the University of Victoria

Vol. 26

Issue 1 (September 2016)

# Working Papers of the Linguistics Circle of the University of Victoria

## – Vol. 26 –

Published by the graduate students
of the University of Victoria Linguistics Department

Department of Linguistics
University of Victoria
P.O. Box 3045
Victoria, B.C., Canada
V8W 3P4

ISSN 1200-3344 (print)
ISSN 1920-440X (digital)

http://journals.uvic.ca/index.php/WPLC | wplc@uvic.ca

# Table of Contents

## Acknowledgments

We would like to acknowledge and thank the authors for their contributions that allow us to have such variation in our publication this year. We would also like to thank Alex D'Arcy, who has supported and mentored us in this year's journey. In addition, we would like to thank our committee: Oluwabukola Ariyo, Emily Comeau, Ildara Enríquez García, Nicole Hildebrand-Edgar, and Silas Romig. WPLC would not be possible without their dedicated work. A special thanks is in order to our dedicated copy editor Jasmine El Gazzar. We would also like to thank the UVic Work Study program for their continued support of WPLC, which allows us to hire an undergraduate student to aid with the production of the journal. The digital form of WPLC is supported by Inba Kehoe and UVic Libraries Office of Scholarly Communications.

## Preface to the 26th volume

Welcome to the 26th volume of the Working Papers of the Linguistics Circle!

This current volume of WPLC is a continuation of the tradition at the University of Victoria to provide opportunities for linguistics students to publish their research, both from UVic and other universities. For the past few years, WPLC has been dedicated to the publication of specialized volumes, such as our volume on minority languages or the most recent conference proceedings of the North West Linguistics Conference. This year we have decided to open the field for all submissions, and as a result we are proud to present a collection of current papers that exhibit the wide range of research areas that linguistics students pursue.

The papers in the collection come from various disciplines. We open this year's volume with a sociolinguistics study on the local English spoken in Victoria, continue with two studies on second language acquisition, followed by a linguistic analysis of variation in historical corpus, and finish with a comparative phonetics study on consonant clusters.

We hope that this variety of topics can not only benefit our readers in their specific areas of interest, but that it also serves as an invitation to current and future students to publish their research and exchange ideas with a community of graduate students.

Editorial Committee,
WPLC 26

# Editorial committee

# Disentangling frequency effects and grammaticalization

**Nicole Hildebrand-Edgar**
University of Victoria
*nchil@uvic.ca*

This paper presents a case study of I DON'T KNOW to investigate the nature of the phonetic reduction using data from the Victorian English Archive (D'Arcy 2011-2014). This phrase has a high usage frequency and is commonly reduced in speech, two concomitant processes in grammaticalization. Further, I DON'T KNOW has use beyond its referential function of "lack of knowledge": it serves various pragmatic functions as a discourse marker. The relationship between phonetic form and semantic function is investigated using quantitative variationist analysis. Similar patterning to that previously reported for other varieties of English is found, suggesting a universal pathway of grammaticalization. Moreover, the frequency of phonetically reduced pragmatic tokens increases in apparent time, suggesting ongoing change of the discourse marker. This change is argued to constitute ongoing phonetic reduction of an already grammaticalized form. The role of frequency effects in driving ongoing change is restricted to one form while the semantic functions remain stable.
*Keywords: Grammaticalization; frequency effects; variationist sociolinguistics; language variation and change*

## 1 Introduction

In the course of everyday speech, common phrases are often subject to phonetic erosion, or the reduction of speech sounds. This includes such commonly heard forms as *whatcha doing* (in lieu of *what are you doing*), *hafta* (for *have to*), and many others. A particularly salient example is *I dunno,* a phonetically reduced form of I DON'T KNOW[1] (first noted in Kaisse, 1985). I DON'T KNOW may be reduced to such a degree that it surfaces as a prosodic grunt, only identifiable by its rise-fall intonation. The fact that *I dunno* appears orthographically in text and Internet correspondence is further testament to its prevalence in the minds of speakers. The alternation between the full and reduced variants of I DON'T KNOW does not appear to be conditioned by any linguistic factor; rather, Scheibman (2000) suggests the variation between forms is subject to speaker choice. On the assumption that seemingly free variation between variants is rarely, if ever, random (Labov, 1969), a Variationist Sociolinguistic approach is

[1] The surface form is represented by italics, and the underlying form by capital letters.

adopted to determine the relationship between the full and reduced variants of I DON'T KNOW and the internal and external factors that influence this alternation. In particular, this analysis aims to determine whether grammaticalization is implicated in the variation.

Quantitative studies have shown that phonological reduction is linked to token frequency within the context of grammaticalization (e.g. Thomson & Mulac, 1991), and that the frequency of a collocation drives the propensity for contraction or reduction (Lorenz, 2013). Bybee (2006) outlines the process of reduction in the context of grammaticalization: widening contexts of use and increased frequency lead to entrenchment of collocations into single processing units (or "chunks"); subsequently, they are accessed and produced with less effort, and are thus subject to phonetic reduction. In a corresponding process, the loss of semantic content and prosodic weight may cause loss of stress, promoting reduction. This proposed two-part process is what Bybee, Perkins, & Pagliuca (1994) term the *Parallel Reduction Hypothesis*. The reduction that ensues affects both articulatory gestures and temporal durations. Consequently, an erstwhile multi-morphemic phrase (such as *go + ing to*) comes to be fused and compressed (*gonna)* (Bybee et al. p. 6). Lorenz argues that this "parallel" process is more accurately described as cyclical: as forms progress through stages of grammaticalization, desemanticization (loss of semantic content) leads to phonetic reduction, but phonetic reduction does not lead to further desemanticization. He proposes that the Parallel Reduction Hypothesis be rephrased as a cycle of reanalysis leading to reduction.

Thus, to show that a form is commonly reduced does not, in itself, imply that the form has grammaticalized. As argued in Hopper (1991), while reduction processes such as condensation (shortening of forms) and coalescence (collapsing of adjacent forms) typically accompany grammaticalization, they are neither necessary nor sufficient diagnostics. Rather, they are typical of forms that are advanced in terms of grammaticalization. Forms that are not grammaticalized may also be reduced. The challenge for identifying instances of grammaticalization in synchronic studies is disentangling general frequency effects and the frequency-driven changes associated with grammaticalization: semantic fading (or bleaching), phonological reduction, positional fixing (or syntactic rigidity), and erasure of word boundaries (Bybee et al., 1994; Hopper & Traugott, 2003).

Several of these changes are demonstrated in conversational use of I DON'T KNOW. In addition to its referential use in expressing insufficient knowledge, studies have shown that I DON'T KNOW is deployed for pragmatic functions, including turn-management, hedging, politeness, and face-saving (Baumgarten & House, 2010; Beach & Metzger, 1997; Bybee & Scheibman, 1999; Pichler, 2009; Diani, 2004; Weatherall, 2011). Following is an example of I DON'T KNOW as a politeness device provided in Grant (2010, p.2286) from the CANCODE corpus:

[At a travel agent's]
S1: Did you want to take out insurance?
S2: Erm I'd like to ask about it but I don't know if I want to do that today.
S1: Okay. (CANCODE)

Here, I DON'T KNOW serves to soften S2's refusal of insurance, thereby protecting the face of S1. Both Scheibman (2000) and Pichler (2009) report strong correlations between the full form and the referential function—expressing lack of speaker knowledge—and between the reduced form and pragmatic functions. This form-function split is attributed to grammaticalization of the construction.

The question that arises is whether the findings reported in Scheibman (2000) and Pichler (2009) for I DON'T KNOW are due to a universal path of grammaticalization based on its semantic function, or whether this variation is conditioned by community-specific factors. To address this question, this project investigates the functional and social conditioning of I DON'T KNOW in a corpus of speech data from Victoria, BC. If the findings reported in Scheibman (2000) and Pichler (2009) are indeed indicative of a universal path of grammaticalization, similar results are predicted in this different variety of English. This exploratory analysis will contribute to the literature concerning the grammaticalization of constructions, and will have implications for analyzing the complex relationship between frequency, phonetic reduction and grammaticalization.

## 2 Literature Review

### 2.1 Frequency and reduction in grammaticalization

Though Meillet's (1912) original conceptualization of grammaticalization applied to single word-forms, recent work has shown that constructions may also be grammaticalized (e.g. Bybee, 2006; Torres Cacoullos & Walker 2009a, Lorenz 2013). Thompson & Mulac (1991) make a case for grammaticalization that extends beyond individual lexemes in an analysis of *that*-deletion and epistemic parentheticals in English. They find that the most frequent subject-verb combinations without *that* occur most frequently as epistemic parentheticals, the verbs encoding subjective meanings associated with belief and mode of knowing. Their findings indicate that grammaticalization is reliant on discourse frequencies and recurrent patterns.

The notion that grammar arises from one's experience with patterns of language is foundational to usage-based models of language. From this framework, Bybee (2006) investigates the role of frequency in grammaticalization of constructions. She provides evidence that frequency is an important factor in grammaticalization, as it promotes both the autonomization of new constructions (that is, cognitive independence from their source forms), as well as phonetic reduction of these constructions. When constructions are encountered with increasing levels of frequency, they may become conventionalized (as idioms or prefabrications). With higher frequency, new constructions with their own categories may be established. Extremely high frequency may then lead to grammaticalization of these new constructions and changes in constituency. Bybee

states that certain changes associated with grammaticalization are, in part, conditioned by frequency: autonomy, semantic bleaching (or semantic change), and reanalysis (loss of morphosyntactic boundaries) (pp. 720-721). As constructions are encountered more frequently, they are produced more fluently, and this phonetic reduction accumulates in the cognitive representation. This reduction process is recurrent, as already reduced variants of high-frequency phrases are more often selected for production, and subsequently undergo further reduction. Bybee argues that these frequent constructions are single processing units (or, what she calls chunks), making them susceptible to further reduction and grammaticalization. Therefore, this process of reduction only occurs on the grammaticalized form (p. 724).

An important caveat is that new constructions may arise without grammaticalization: certain general constructions may develop new pragmatic meaning without being completely disassociated from their source meanings (such as *How do you do?,* which is still associated with the source question *What are you doing?*) (Bybee, 2006, p. 723). Lorenz (2013) addresses this issue in regards to the reduced forms *gonna, gotta,* and *wanna.* He asks whether they are simply typical ways of pronouncing *going to, got to,* or *want to*, or whether they have independent meanings from their source forms and distinct cognitive representations. He argues that the contracted forms *gonna, gotta,* and *wanna* are emancipated (autonomous) from their source forms as a result of becoming entrenched in memory through frequent usage. In this process of emancipation, a full form becomes phonetically reduced, and as this reduced form is frequently used, it becomes a conventional expression encoding particular meanings (the process of divergence). As the reduced form becomes its own lexical item, speakers stop interpreting it as the full form, and the initial motivation for reduction is lost.

The phrase I DON'T KNOW is both commonly reduced and extremely frequent in discourse. Investigating both the BNC and COCA corpora, Baumgarten & House (2010) find that I DON'T KNOW is the most frequently occurring negative collocation. Previous analyses also show that I DON'T KNOW is a highly frequent collocate across varieties of English (Kaisse, 1985, Scheibman 2000). This high frequency of usage, in addition to the existence of the reduced for *I dunno,* suggests that I DON'T KNOW may very well be a grammaticalized construction. However, while frequency and reduction are processes that occur within grammaticalization, there must be evidence that an erstwhile lexical (content) form has changed in such a way as to assume characteristics of a grammatical (functional) form in order to validate this claim (Hopper & Traugott, 2003). The following section provides a review of the pragmatic functions that are associated with I DON'T KNOW.

## 2.2 Functional analyses of I don't know

In everyday conversation, I DON'T KNOW is deployed for a much broader range of functions than to simply claim lack of knowledge. Grant (2010), analyzing usage patterns of *I don't know* and *I dunno* in text corpus data of British and New Zealand speech, finds that both *I don't know* and *I dunno* can be used as epistemic devices and expressions of stance (p. 2290). *I dunno* is used especially as a politeness device to soften disagreement. Grant reports that the full form *I don't know* has a greater range of usage, while the reduced form *I dunno* is predominantly a hedging device. While Grant's findings are useful for acknowledging the various functions that *I don't know* and its reduced variant can serve, the form-function regularities she outlines are questionable: as she acknowledges, the data is only written, and there is no way of confirming the criteria by which the transcribers differentiated the two forms.

Weatherall (2011) examines the functional distribution of I DON'T KNOW in British, New Zealand, and American speech. Her analysis focuses specifically on instances of I DON'T KNOW that have scope over the following proposition (as opposed to those that are responses or follow an assessment). She finds that these prepositioned tokens fall into two broad categories: those used in first assessments (signalling exaggeration or non-seriousness), and those used in approximations. In both cases, I DON'T KNOW indexes lack of speaker commitment to what follows. Weatherall argues that these prepositioned epistemic hedges function to disclaim knowledge authority (especially in the first assessment cases), which indicates that source meaning (lack of knowledge) persists. Similarly, Diani (2004) finds that I DON'T KNOW can function to avoid explicit disagreement, avoid commitment, minimize face-threatening acts, and mark uncertainty—all of which retain the central meaning of lack of knowledge. However, neither Weatherall (2011) nor Diani (2004) makes a distinction between full- and reduced-forms of *I don't know*.

The variation between full and reduced forms of I DON'T KNOW is addressed in Scheibman (2000). The linguistic conditioning that drives the reduction of the negative auxiliary *don't* is explored by analyzing its use in everyday conversation. Conversational data from American speakers reveals that reduced DON'T occurs in limited but highly frequent collocations, predominantly I DON'T KNOW. By comparing the semantic and interactive contexts of the full and reduced forms of I DON'T KNOW, Scheibman finds a form-function regularity: both full and reduced forms may express the referential meaning of insufficient knowledge, but only reduced *I dunno* is used in pragmatic (textual or organizational) or subjective (face saving or politeness) functions. These functional correlations are therefore inconsistent with those reported in Grant's (2010) text corpus analysis, where the more full form could serve either referentially or pragmatically, and the reduced form only pragmatically. Scheibman contends that grammaticalization is not implicated for reduced *don't* itself; rather, in the spirit of Thompson & Mulac's (1991) proposal, the conventionalized expressions in which it most frequently occurs are

grammaticalized (e.g. *I don't know, I don't think*). These conventionalized expressions are processed as single units or "chunks", and tend to have pragmatic functions. The full forms continue to exist, though with different functions (layering), and the new forms come to take on subjective and textual meanings (semantic bleaching). Scheibman's results must be taken with caution, however, as they are based on very few tokens (N=36).

Pichler (2009) also addresses the phonetic variants of *don't.* The differential distribution of discourse markers I DON'T KNOW and I DON'T THINK is investigated in everyday speech in Berwick-upon-Tweed. The study employs qualitative methods of conversation analysis in determining functional and social meanings, and quantitative methods of Labovian sociolinguistics in analyzing linguistic conditioning. Pichler identifies three non-localized phonological variants of I DON'T KNOW: a full form *I don't know* (with a marked boundary between the *n* of DON'T and the *n* of KNOW and a full vowel *o* in DON'T), an intermediate form *I donno* (with no marked boundary and a full vowel), and a reduced form *I dunno* (with no marked boundary and a reduced vowel). The reduced form is found to be the most frequent variant across social groups (aside from older males), and has the greatest potential to occur in all pragmatic functions. The full form is found to correlate strongly with referential functions. In addition, a localized variant, *I divn't knaa,* is identified, and found to be socially conditioned. The functional conditioning of non-localized variants is claimed to be a result of grammaticalization, as the distribution of forms exhibits various indices of grammaticalization from Hopper (1991): the full form dominates in referential contexts while an intermediate form is used across functions (layering), the reduced form is very rarely intervened by adverbial modification (decategorialization), and the source meaning of "lack of knowledge" is maintained in the reduced epistemicity meaning of the grammaticalized forms (persistence). As in Scheibman (2000), it is argued that I DON'T KNOW is a formulaic, single processing unit—a fact which has enabled its grammaticalization. Pichler (2009) further suggests that the reduced variable may still be increasing in positional mobility and discourse functions.

### 2.3 Summary of literature review

Frequency and phonetic reduction are inherent in grammaticalization, although they are not, in themselves, sufficient for identifying forms that have grammaticalized. I DON'T KNOW has been found to be a highly frequent collocation across varieties of English, and its propensity for reduction has been noted in multiple studies. It has also been observed in functional analyses that I DON'T KNOW is used in everyday conversation to encode a variety of pragmatic meanings in addition to its referential meaning. Several authors, notably Scheibman (2000) and Pichler (2009), have found a form-meaning relationship for variants of I DON'T KNOW, and attribute this relationship to grammaticalization of the construction. If these findings imply a universal path

of grammaticalization of I DON'T KNOW, similar results should emerge from analyses of its distribution in other English speaking communities.

## 3 Methodology and Data

### 3.1 Theoretical assumptions

To further explore the form-function regularities of full and reduced variants, and to test whether the findings in Scheibman (2000) and Pichler (2009) are due to a universal path of grammaticalization of I DON'T KNOW, I likewise examine its use in natural spoken conversation. As stated in the introduction, a Variationist Sociolinguistic approach is adopted for this analysis. In this framework, the fact that individual speakers will exhibit variable behaviour is recognized; thus, inherent variability in everyday language is taken into account. Further, generally accepted indices of grammaticalization—layering, persistence, semantic bleaching, syntactic generalization, and phonetic erosion—make predictions that can be tested using a variationist approach (Walker, 2010, p. 106). A multivariate analysis is employed using GoldVarbX (Sankoff, Tagliamonte, & Smith, 2012) to tease apart the complex interaction of social and linguistic factors that influence speaker choice. The resultant form-function patterns will be examined in order to determine if grammaticalization is implicated in the variation.

This analysis further assumes the concept of a cline of grammaticalization (Hopper & Traugott, 2003). That is, grammaticalization does not involve abrupt shifts from one category to another, but consists of a series of small transitions that emerge synchronically as a continuum between a fuller, less grammatical form, to a reduced, more grammatical form (p. 6). This assumption becomes important when interpreting results that emerge from the data.

### 3.2 Data and Coding

Data was extracted from the Synchronic Corpus of Victoria English (SCVE), housed at the University of Victoria Sociolinguistics Research Lab (SLRL). The corpus consists of sociolinguistic interviews with 162 speakers from Victoria, BC, born between 1913 and 1996. A total of 24 speakers were selected based on the factors of age and gender (Table 1). In total, this smaller set of interviews comprises 21 hours of speech and over 275 000 words. Three age groups were defined: younger (18-25), middle (30-49), and older (63-85) to enable an apparent-time analysis of the distribution of variants (Tagliamonte & D'Arcy, 2007), that is, an analysis of different generations at one point in time. As this analysis does not address localized variants or prestige forms, socioeconomic status was not included as an independent variable; all speakers have mid- or upper mid-range SES scores.

| | Male | Female | N Speakers | N Tokens |
|---|---|---|---|---|
| **17-25** | 4 | 4 | 8 | 111 |
| **30-49** | 4 | 4 | 8 | 91 |
| **63-85** | 4 | 4 | 8 | 83 |
| **TOTAL** | 12 | 12 | 24 | 285 |

*Table 1: speaker sample and tokens extracted from SCVE*

All instances of the negative periphrastic DO in collocation with the verb *know* and the first-person pronominal subject *I* (or a zero-subject that is coreferential with *I*) were extracted from a 30-minute segment of each interview. On the assumption that speech is more monitored at the beginning of an interview, the 30-minute segment after first 10 minutes of the interview was used for analysis. To control for vast divergences in rates of usage that may confound the results, this 30-minute window was decreased or increased so that speakers had no more than 20 tokens or no fewer than 5 tokens. Where the form or function could not be unambiguously determined, such as in utterances that were cut off or obscured by other sounds, tokens were excluded. Twelve tokens were also excluded because they included adverbial modification (e.g. *I really don't know, I don't even know*). These forms will be addressed in the discussion. Following these methods, a total of 185 tokens were retained and coded for social and linguistic factors.

Each token of *I don't know* was coded auditorily for phonetic form. The full form (*I don't know*) has a distinct morpheme boundary (normally a glottal stop) between the nasals of *don't* and *know*, and the full vowel [o] in *don't*. The intermediate form (*I donno*) has no distinct morpheme boundary between *don't* and *know*, but still has a full vowel in *don't.* The reduced form (*I dunno*) has no distinct morpheme boundary between *don't* and *know*, and the vowel in *don't* is reduced to [ə]. These three forms are similar to those identified in Pichler (2009). The present analysis also includes a category for further reduced forms (*I d'no*), which have no morpheme boundary between *don't* and *know*, a reduced vowel, and some further reduced aspect, such as a lenited [d] in *don't* (e.g. [əno]), no vowel at all in *don't* (e.g. [dno]), a complete fusion of *don't* and *know* (e.g. [ɾo] with a flap), or a complete lack of phrase-medial consonants (observed as a 'prosodic grunt' with a rise-fall intonation that identifies it as I DON'T KNOW).

Following Pichler (2009), syntactic configuration was coded by determining whether tokens have an overt complement. Bound tokens are either preceded or followed by an overt complement, as in (1) and (2) respectively:

(1) WB/79/f well <u>what became of him</u> **I don't know** but I suppose he'd have been relocated

(2) JS/23/m **I don't know** <u>about weddings and stuff</u>

Unbound tokens have no overt complement and are grammatically independent, as in (3):

(3) KA/18/m my most embarrassing moment **I don't know** . I don't think that's a very good question for me

Semantic function was determined following various observations in the literature regarding the pragmatic functions of I DON'T KNOW (e.g. Baumgarten & House, 2010; Beach & Metzger, 1997; Bybee & Scheibman, 1999; Diani, 2004; Weatherall, 2011), and noting indications from the prosody, conversational context, and occurrence of other discourse markers (Pichler, 2009). Tokens that indicated a lack of knowledge were coded as Referential, as in (4) and (5):

(4) CA/21/f so there was actually like a T-A at the school who would take me outside . on my bike and show me how to ride a bike **I don't know** why I got this weird special treatment

(5) JF/84/m so I've been there eleven years . and I applied for Quadra why I did **I don't know**

I DON'T KNOW is also used to maintain rapport and mitigate face threats. Tokens that functioned as markers of reduced epistemicity, politeness devices, or hedges in communicating lack of commitment to a following or preceding utterance were coded as Interpersonal, as in (6) and (7):

(6) INT: would you . put birds in there ? are you interested in doing that ?
BD/30/m uh . no not really it seems **I don't know** it seems weird to . [INT:laughs] keep them in a cage for your viewing . when they could just fly around . so <yeah> . [clicks tongue] yeah

(7) INT how did your pajama pants turn out ?
BL/31/f they were great but . um . pajama pants **I don't know** . they're not that special [laughs]

I DON'T KNOW may be deployed to structure dialogue. It is available to mark topic boundaries, initiate or prevent turn exchange, and link aborted and recast statements (repair). These tokens were coded as Textual, as in (8) and (9):

(8) BL/31/f she's good she was always around . <yeah> because she um . yeah she stayed at home she had a few jobs occasionally but <right> mostly she was at home <mm-hm> .um yeah had a good childhood . <yeah> um . we . you know **. I don't know**

(9) DK/63/m we'd jump down the laundry shoots to land on the mattresses
INT that's /awesome/
DK/63/m /**I don't know** / and then they had those dumb {unclear} we used to crawl up that but uh .

Finally, as Pichler (2009) notes, discourse markers are polypragmatic devices. In the case of I DON'T KNOW, many occurrences serve both interpersonal and textual functions. Rather than subjectively choosing between Interpersonal and Textual, these tokens were coded as Polypragmatic, as in (10) and (11):

(10) *Repair (textual) and hedge (interpersonal)*
CA/21/f she'd talk about how . **I dunno** she'd she made a lot of . World-War-Two jokes [laughs] . i-- with the understanding that it was a terrible thing but it you know . {unclear} . you could make light of it

(11) *Turn-yield (textual) and disclaimer (interpersonal)*
JF/84/m um I worked at Macaroni-Grill on Davie that was the {unclear} Mansion ? <oh okay> **I don't know** yeah

In coding tokens in such a manner the effects of age, gender, syntactic form, and semantic function can be quantified and statistically analyzed. This enables a statistical analysis of the internal and external factors that affect variant choice.

## 4 Results

### 4.1 Overall frequency

To ascertain the frequency of usage of I DON'T KNOW, the entirety of the 18 interviews selected from the corpus were analyzed using AntConc concordance software (Anthony, 2011). Consistent with results reported in previous corpus analyses (E.G. Baumgarten & House, 2010; Grant, 2010), I DON'T KNOW is the most frequent 3-word phrase in these materials, occurring a total of 707 times. The next most frequent 3-word phrase, *a lot of*, occurs 447 times.

Following Thompson and Mulac (1991), type frequency of I DON'T KNOW was compared to the token frequency of the negative periphrastic DO

construction. Negative periphrastic DO occurs a total of 1545 times, meaning that nearly half of its occurrences are in the construction I DON'T KNOW (707/1545 = 46%). As shown in Table 2, the first person singular pronoun *I* is by far the most frequent subject collocating with DON'T (1153/1545= 75%). Table 3 shows that when the following word is taken into account, *know* is the most frequent collocate (748/1545=48%).

| | N | % |
|---|---|---|
| ***I + don't*** | **1153** | **75** |
| ***you + don't*** | 107 | 7 |
| ***they + don't*** | 59 | 4 |
| ***we + don't*** | 36 | 2 |
| ***OTHER*** | 190 | 12 |
| **TOTAL** | 1545 | |

*Table 2: relative frequencies of subjects collocating with negative periphrastic DO*

| | N | % |
|---|---|---|
| ***don't know*** | **748** | **48** |
| ***don't think*** | 124 | 8 |
| ***don't have*** | 63 | 4 |
| ***don't like*** | 49 | 3 |
| ***don't want*** | 46 | 3 |
| ***don't remember*** | 45 | 3 |
| ***OTHER*** | 470 | 30 |
| **TOTAL** | 1545 | |

*Table 3: relative frequency of verbs following negative periphrastic DO*

Considering the findings regarding frequency and reduction from Thompson & Mulac (1991) cited in §2, the high token frequency of I DON'T KNOW is conceivably a major cause of its phonetic reduction. Whether or not this frequency has led to phonetic reduction in the context of the emergence of grammar will be determined by analyzing the social and linguistic conditioning of the variants. If a connection is established between phonetic form and function, this will indicate that the phonetic reduction is not simply an effect of frequency, and will support the grammaticalization hypothesis.

As outlined in §3, 268 tokens were coded for 3 linguistic factors: phonetic form, syntactic form, and semantic function. Cross tabulation revealed that the factor groups Syntactic form and Semantic function strongly interact: nearly categorically, syntactically bound tokens (those with an overt complement) were used in referential contexts (88/93=95%), and syntactically unbound in other pragmatic functions (163/174=94%). Bound tokens were therefore excluded from the analysis, leaving a total of 174 unbound tokens. This

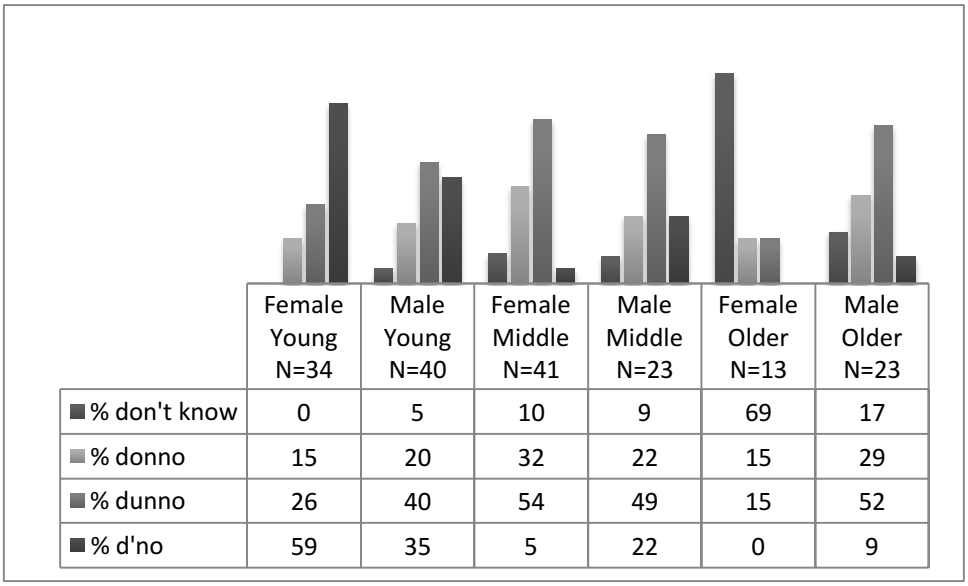
resulted in excluding over half of *don't know* and *dunno* tokens, as well as every instance where *dunno* and *d'no* were used referentially. This will be further discussed in §5. Table 4 shows the overall distribution of variants after bound tokens were removed.

| | ***don't know*** | ***donno*** | ***dunno*** | ***d'no*** | **TOTAL** |
|---|---|---|---|---|---|
| **%** | 12 | 22 | 41 | 25 | |
| **N** | 21 | 38 | 72 | 43 | 174 |

*Table 4: overall distribution of variants*

### 4.2 Social factors contributing to distribution of variants

Figure 1 shows the distribution of variants with respect to age and gender. The figure shows the percentage of each of the variants in each of the 6 social groups: young female, young male, middle female, middle male, old female, old male. While Pichler (2009) finds that the reduced form *I dunno* is the most common variant across social cohorts (aside from old male), the present data reveal a different pattern: the full form *I don't know* is the most common variant in the old female group, the reduced form *I dunno* is favoured by both middle aged groups and older males, and the further reduced form *I d'no* is the most favoured in the young cohorts. The intermediate form *donno* is not conditioned by age, but does correlate with male speakers. This may contribute to the difference between the old male and old female cohort; cross tabulating the results did not reveal any other factor that may affect this difference, although with such small numbers (N=13 for old female and N=23 for old male), ideolectal effects could easily obscure the results.



| | Female Young N=34 | Male Young N=40 | Female Middle N=41 | Male Middle N=23 | Female Older N=13 | Male Older N=23 |
|---|---|---|---|---|---|---|
| % don't know | 0 | 5 | 10 | 9 | 69 | 17 |
| % donno | 15 | 20 | 32 | 22 | 15 | 29 |
| % dunno | 26 | 40 | 54 | 49 | 15 | 52 |
| % d'no | 59 | 35 | 5 | 22 | 0 | 9 |

*Figure 1: social distribution of variants of I DON'T KNOW.*

This age effect shown in figure 1 is an unexpected result given previous analyses of I DON'T KNOW. The incrementally increasing frequency of the very reduced form and the decreasing frequency of the full form in apparent time, as observed in figure 2, suggests that this reduction is not an age-graded effect, but ongoing generational change (Tagliamonte & D'Arcy, 2007). However, the only decisive method for resolving this would a study in real time. Determining whether this change has to do with phonetic reduction only, or whether this is an instance of the emergence of grammar, requires an analysis of the functional distribution of forms.

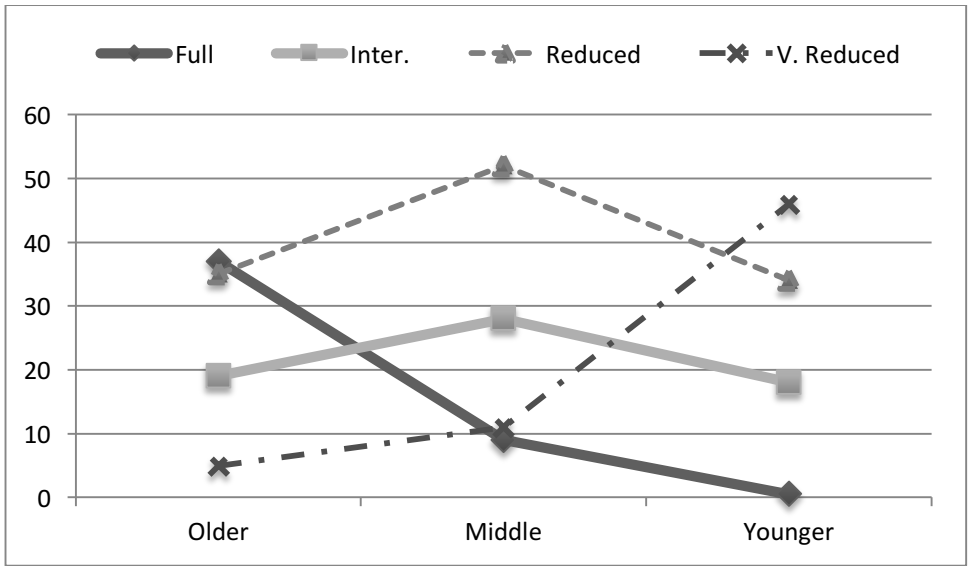
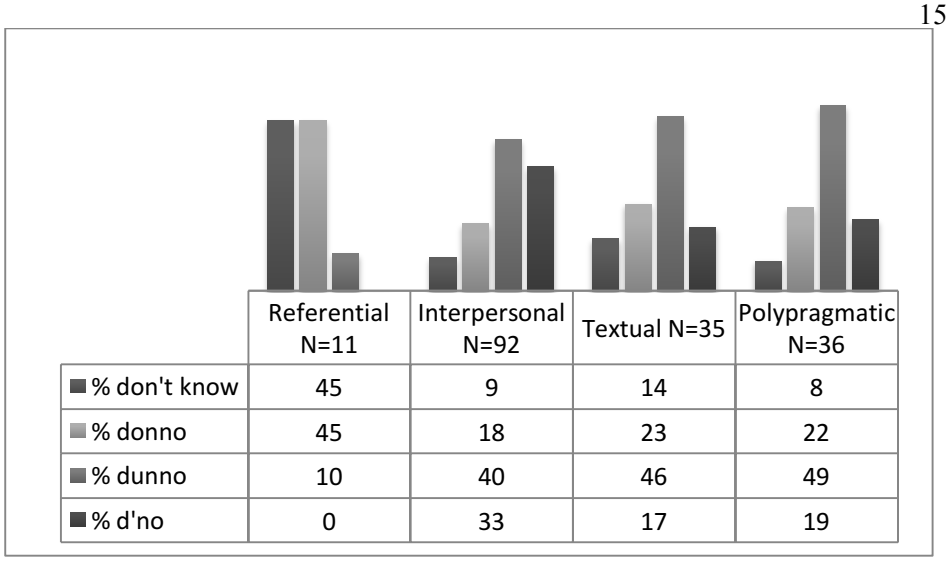



*Figure 2: age distribution of variants of I DON'T KNOW*

### 4.3 Linguistic factors contributing to the distribution of variants

Figure 3 gives a breakdown of the functional distribution of the variants of I DON'T KNOW. It tracks the percentages of each of the four categories (referential, interpersonal, textual, and interpersonal-textual) across the four variants. The patterning here is similar to that reported in Scheibman (2000) and Pichler (2009): the full variant correlates with referential uses, though this result must be treated cautiously due to the low number of referential tokens. Recall that, in removing the bound tokens, the majority of tokens functioning referentially, including all instances of *I dunno* and *I d'no* that function referentially, were also removed. The remaining 8 tokens that function referentially are *I don't know* or *I donno*. The reduced variants *I dunno* and *I d'no* correlate with pragmatic uses. These results are suggestive of functional conditioning of variants of I DON'T KNOW as a result of the variable's grammaticalization. The patterning in figure 3 also shows that the pragmatic functions—interpersonal, textual, and polypragmatic—are not particularly differentiated, as they all pattern in the same way. For this reason, they will be collapsed into a singular pragmatic category for the following analysis.



| | Referential N=11 | Interpersonal N=92 | Textual N=35 | Polypragmatic N=36 |
|---|---|---|---|---|
| % don't know | 45 | 9 | 14 | 8 |
| % donno | 45 | 18 | 23 | 22 |
| % dunno | 10 | 40 | 46 | 49 |
| % d'no | 0 | 33 | 17 | 19 |

*Figure 3: functional distribution of Variants of I DON'T KNOW.*

### 4.4 Multivariate analysis of the distribution of variants

The results in §4.2 suggest that the distribution of variants seems to be shifting in apparent time: older speakers are more likely to use the full form, middle speakers the reduced form, and younger speakers the very reduced form. In §3.3, the results indicate that reduced forms, and primarily *dunno*, correlate with pragmatic functions. To test whether the apparent ongoing change in phonetic reduction is an effect of differential uses of semantic function across age groups, a multivariate analysis is performed using GoldVarbX. Referential function was included only for the analyses of the full and intermediate forms (Tables 5 and 6). As none of the 8 referential tokens were reduced or very reduced variants, they were excluded from the multivariate analysis of these tokens (Table 7 and 8). [2] During the initial analysis, it was noted that one speaker categorically produced the very reduced form *I d'no.* She contributed nearly half (18/43=42%) of all *I d'no* tokens. Further, this speaker was in the young cohort. As the total number of tokens is only 174, such a speaker effect can have a huge difference on the distributional results. The data from this speaker was excluded, and the remaining 156 tokens were rerun through GoldVarb.

[2] Investiation of these 8 referntial tokens revealed no speaker effects: the 8 tokens came from 7 different speakers, spread across age groups (2 young speakers, 2 middle, and 4 older) and speaker sex (4 female speakers and and 4 male).

| | | | |
|---|---|---|---|
| | | ***don't know*** | |
| **Input** | | .081 | |
| **Overall %** | | 14 | |
| **Total N** | | 156 | |
| | **FW** | **%** | **N** |
| **Sex** | | | |
| Female | **.65** | 17 | 70 |
| Male | **.37** | 9 | 86 |
| *Range* | ***28*** | | |
| **Age** | | | |
| Older | **.86** | 36 | 36 |
| Middle | **.43** | 9 | 64 |
| Younger | **.30** | 3 | 56 |
| *Range* | ***56*** | | |
| **Semantic Function** | | | |
| Referential | **.86** | 46 | 11 |
| Pragmatic | **.47** | 10 | 145 |
| *Range* | ***39*** | | |

*Table 5: multivariate analyses of the contribution of internal and external predictors (significant and non significant) to the probability of full form.*
(Log likelihood= -46.86, p=0.044)

| | | | |
|---|---|---|---|
| | | ***donno*** | |
| **Input** | | .237 | |
| **Overall %** | | 24 | |
| **Total N** | | 156 | |
| | **FW** | **%** | **N** |
| **Sex** | | | |
| Female | [.55] | 29 | 70 |
| Male | [.46] | 21 | 86 |
| **Age** | | | |
| Middle | [.54] | 28 | 64 |
| Younger | [.51] | 23 | 56 |
| Older | [.42] | 19 | 36 |
| **Semantic Function** | | | |
| Referential | [.74] | 46 | 11 |
| Pragmatic | [.48] | 23 | 145 |

*Table 6: multivariate analyses of the contribution of internal and external predictors to the probability of intermediate form.*

| | | | |
|---|---|---|---|
| | | ***dunno*** | |
| **Input** | | .045 | |
| **Overall %** | | 46 | |
| **Total N** | | 156 | |
| | **FW** | **%** | **N** |
| **Sex** | | | |
| Female | [.50] | 47 | 70 |
| Male | [.50] | 45 | 86 |
| **Age** | | | |
| Middle | [.56] | 52 | 64 |
| Younger | [.48] | 45 | 56 |
| Older | [.44] | 39 | 36 |
| **Semantic Function** | | | |
| Pragmatic | **.54** | 49 | 145 |
| Referential | **.11** | 9 | 11 |
| *Range* | ***43*** | | |

*Table 7: multivariate analyses of the contribution of internal and external predictors (significant and non significant) to the probability of reduced form* (Log likelihood = -114.98, p= 0.015)

| | | | |
|---|---|---|---|
| | | ***d'no*** | |
| **Input** | | .12 | |
| **Overall %** | | 16 | |
| **Total N** | | 156 | |
| | **FW** | **%** | **N** |
| **Sex** | | | |
| Male | **.67** | 24 | 86 |
| Female | **.30** | 5 | 70 |
| *Range* | ***37*** | | |
| **Age** | | | |
| Younger | **.67** | 29 | 56 |
| Middle | **.50** | 11 | 64 |
| Older | **.24** | 6 | 36 |
| ***Range*** | ***33*** | | |

*Table 8: multivariate analyses of the contribution of internal and external predictors (significant and non significant) to the probability of very reduced form (Log likelihood = -78.25, p= 0.000).*

The results from table 5 show that for the full form, variant choice is significantly favoured for referential uses, and among speakers who are female and in the older category. In table 6, no factors reach significance for the intermediate form, though there is a robust effect of referential function. Table 7 shows that for the reduced form, semantic function is the only factor selected as significant, and the variable is favoured for pragmatic uses. There is also a direction of effect that privileges the middle age cohort. Finally, table 8 shows that age has a very strong effect for the very reduced form, with younger speakers being the most likely to use it. This variant is also more favoured among male speakers, though this did not reach significance. The very reduced form was categorically pragmatic, so the semantic factor group is not included in the analysis.

In Pichler (2009), the differential distribution of the full and reduced form is found to be significant across referential uses and pragmatic uses *combined.* Because there were no referential tokens that surfaced as reduced or very reduced variants in the present data, a similar comparison could not be made. An second analysis was performed which included syntactic function as a factor group instead of semantic function, but this did not result in a better fit for the data. Further, it resulted in a similar problem: removing all referential tokens from the analysis also removed the majority of bound tokens.

## 5 Discussion

In analyzing these variants, it must be noted that the four categories identified are not well-defined, immutable groups: they represent different points on a continuum, from full articulation to nearly complete erosion (e.g. a consonantless prosodic grunt). This is a key characteristic of changes that occur in the context of grammaticalization: small shifts occur on a cline from a more full, less grammatical form to less full, more grammatical form (Hopper & Traugott 2003). Admittedly, gradual change is difficult to differentiate from abrupt change in an apparent-time analysis (Walker, 2010). Further, as addressed in Lorenz (2013), evidence is needed to show that reduced forms are not simply easier ways of pronouncing lexical items, but that they have independent meanings and a distinct cognitive representation.

### 5.1 Grammaticalization

Unlike Scheibman (2000) and Pichler (2009), the present results do not yield a phonetic form-function split. However, a split was identified in syntactic context and function, which complicated comparison of the distribution of phonetic variants. As stated in §3.1, bound tokens are nearly categorically used for a referential function (94%), and unbound tokens for pragmatic functions (93%). Figure 4 shows the distribution of variants in these bound tokens. The majority of the excluded bound tokens are the full form *I don't know*.

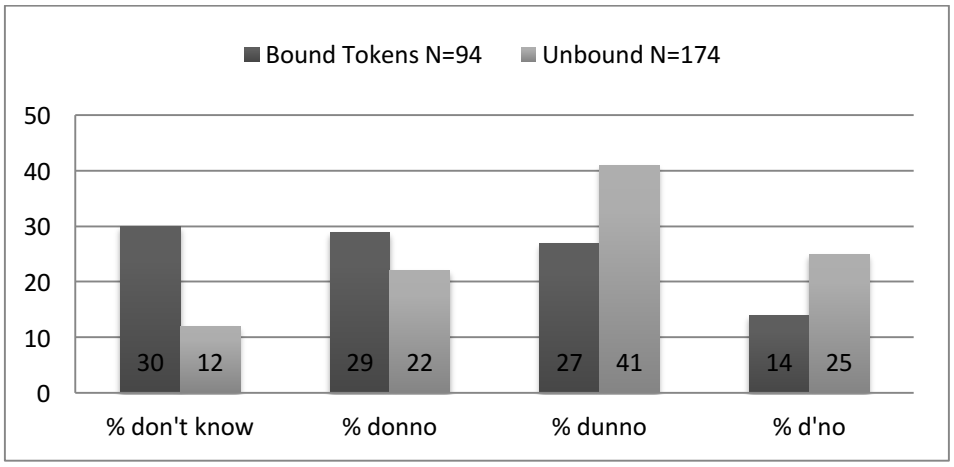



*Figure 4: distribution of variants across bound tokens (N=94) compared to unbound tokens (N=174). The distribution of variants across bound and unbound functions is highly significant ($\chi_2$=22.80, df=3, p<0.01)*

This split indicates two different kinds of I DON'T KNOW: one form encodes the referential meaning of "lack of knowledge" and resides predominantly in matrix-complement (bound) constructions[3], and another that encodes pragmatic meaning and occurs mostly in unbound constructions. This suggests that at one point, the frequent construction I DON'T KNOW spread to wider syntactic contexts, and its use in these context came to be associated with semantically faded, pragmatic meanings. These observations, along with the phonetic erosion of I DON'T KNOW and its high frequency (described in §3.1) are indicative of grammaticalization. I would argue that the generational change observed in the patterning of phonetic variants is not indicative of an emerging new form, but rather the ongoing phonetic reduction of an already grammaticalized form.

Furthermore, several of Hopper's (1991) indices of grammaticalization are observed in the distribution of I DON'T KNOW. Decategorialization is indicated by the fact that the grammaticalized form is rarely used in the matrix-complement construction. Divergence is also observed, as the original lexical form is still comprised of autonomous elements: it may be adverbially modified (e.g. *I don't* ***even*** *know*, *I* ***really*** *don't know)*, while the grammaticalized form rarely has an intervening adverb (Pichler, 2009). Of the 12 adverbially modified tokens that were removed from the analysis, 11 had referential meaning. Persistence is shown in the pragmatic meanings of reduced epistemicity and lack of commitment; nuances of the referential meaning of lack of knowledge remain. Finally, layering is observed in the coexistence of the two forms.

[3] When I DON'T KNOW is in second position as a referential response to a question (as opposed to a polite response, Diani, 2004), there may be no overt surface complement, but a complement is implied, e.g. "I don't know (the answer to the question)"

The finding that phonetic reduction of the grammaticalized form appears to be advancing in apparent time reflects Bybee et al.'s (1994) Parallel Reduction Hypothesis described in the introduction. I DON'T KNOW, as a result of high frequency and widened syntactic context, has become entrenched as a single processing unit autonomous of its source form. Thus, it is accessed and produced faster, and morphological boundaries lose significance and are subject to erosion. The loss of semantic content has enabled I DON'T KNOW to function as a pragmatic marker, which leads to loss of prosodic weight and stress, further catalyzing reduction.

### 5.2 Speaker effects

Because of the small size of my sample, I had to consider how speaker effects complicated my analysis. While some speakers took over an hour to produce 5 tokens, one speaker in particular produced 20 in less than 15 minutes. Not only were tokens highly frequent in her speech, they were nearly categorically very reduced—almost half of the very reduced tokens came from this speaker. This speaker is a female, and is also the youngest in the sample. In initial analyses, when this speaker was included, sex was never selected as significant for any variant. After she was removed, not only did sex emerge as a significant factor for the full form, the direction of effect for the very reduced form switched from female to male.

Though excluded from statistical analyses, qualitative consideration of data from this speaker does give weight to the parallel reduction hypothesis: the only referential tokens in her data were adverbially modified, which indicates that grammaticalization is at a very advanced stage. Adverbial modification is now required to encode referential meaning, and the grammaticalized form is becoming increasingly reduced. Whether this speaker's patterns indicate ongoing change, or linguistic marketplace effects in the heterosexual talk market of high school (Eckert, 2011), or the speaker's own idiolect remains an open question, but an interesting avenue for further research.

### 5.3 Moving forward

This project is an exploratory investigation of variation that potentially indicates grammaticalization. As such, it has brought about many questions and directions for future research. The results reported here are not entirely consistent with those reported by Scheibman (2000) or Pichler (2009). While there are similar correlations observed between the full-form and referential uses and the reduced form and pragmatic uses, the affect of age was not reported in either study. This indicates that something different is happening to I DON'T KNOW in Victoria English—is this a community effect, or is this a stage on the universal pathway of grammaticalization of I DON'T KNOW that was undetected by Scheibman or Pichler? To answer this question requires more data from a wider range of speech communities. It was further reasoned that this increased reduction was a

generational change, not an age-graded effect, and that the grammaticalization of I DON'T KNOW has already taken place. These hypotheses can only be verified by a study in real-time.

The relatively low number of tokens created problems for analysis, especially as no reliability test was performed for the coding procedures. Observed patterns are easily obscured by small differences in raw numbers. Additionally, the small number of speakers and tokens makes it difficult to ensure that patterns observed are not due to ideolectal features. Further, to ensure that the data are representative of individual patterns of usage, Labov (1966, p. 181) advocates for 10-20 instances per speaker, while others call for even more (e.g. Guy, 1993). Several of the speakers in the sample had fewer than 10 tokens. Future analyses will benefit from analyzing a greater number of tokens from a greater number of speakers.

Reduction was defined primarily in terms articulatory gestures for this project. However, Bybee et al. (1994) indicate that phonetic reduction occurs in articulatory gestures as well as duration, though measuring duration was beyond the scope of this project. Further, the unit of focus in coding for reduction was DON'T, following Scheibman (2000) and Pichler (2009). Reduction was most salient for the consonant [d] in DON'T, the morpheme boundary [n?n] between DON'T and KNOW, and the vowel in DON'T. Impressionistically, the initial vowel [aj] ("I") and the final vowel [ow] in KNOW also vary in their phonetic form, though this was not included in coding. The pattern of reduction for these vowels may have interesting implications for the concept of phonetic reduction: does a hierarchy exist for which segments are reduced first? Analyzing variants based on the production of all segments and for phonetic duration would enrich this analysis, potentially yielding interesting results for patterns of distribution.

## 6 Conclusion

The interaction of frequency and phonetic reduction is a well-documented phenomenon in language change (Bybee, 2006). While these two phenomena are inevitable components of grammaticalization, they are not, in themselves, necessary or sufficient in identifying cases of grammaticalization. This study sought to untangle these interwoven processes by performing a variationist analysis of the distribution of phonetic variants of the highly frequent collocation I DON'T KNOW. Results similar to those reported in previous studies were uncovered concerning the form-function regularities of I DON'T KNOW, indicating a universal path of grammaticalization of this construction. Previously unreported results were also found: the increased frequency of reduced and very reduced variants in apparent time indicates ongoing change in the reduction of I DON'T KNOW. Whether this increasing reduction is an age-graded effect or a generational change, and whether it is a community-specific or universal tendency, calls for further examination of this form across time and speech communities.

**Acknowledgements**

A huge thank-you to Alex D'Arcy for her generous expertise and guidance through the many iterations of this project, and to my partner John Edgar for being my ever-patient statistics consultant.

**References**

Anthony, L. (2011). AntConc (Version 3.2.2) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved November 27, 2014, from http://www.laurenceanthony.net/

Beach, W.A., & Metzger, T. (1997). Claiming insufficient knowledge. *Human Communication Research, 23*(4), 562-588.

Baumgarten, N., & House, J. (2010). I think and I don't know in English as lingua franca and native English discourse. *Journal of Pragmatics*, *42*(5), 1184-1200.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*(4), 711-733.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* Chicago, IL: University of Chicago Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of *Don't* in English. *Linguistics*, *37*(4), 575–596.

D'Arcy, A. (2011-2014). Victoria English: its development and current state. Standard Research Grant no. 410-2011-0219. Social Sciences and Humanities Research Council of Canada.

Diani, G. (2004). The discourse functions of *I don't know*. In K. Aijmer & A. Stenström (Eds.): *Discourse patterns in spoken and written corpora*. Amsterdam: John Benjamins. 157-171.

Eckert, P. (2011). Language and power in the preadolescent heterosexual market. *American Speech, 86*(1), 85-97.

Grant, L. (2010). A corpus comparison of the use of I don't know by British and New Zealand speakers. *Journal of Pragmatics*, *42*(8), 2282-2296.

Guy, G. (1993). The quantitative analysis of linguistic variation. In D. Preston (Ed.), *American dialect research*. Philadelphia, PA: John Benjamins. 223-249.

Hopper, P. (1991). On some principles of grammaticization. In E.C. Traugott & B. Heine (Eds.) *Approaches to grammaticalization: Focus on theoretical and methodological issues*, vol. 1. Philadelphia, PA: John Benjamins. 17-35.

Hopper, P. and Traugott, E. (2003). *Grammaticalization.* 2nd edition. Cambridge: Cambridge University Press.

Kaisse, E. (1985). Connected speech: The interaction of syntax and phonology. San Diego, CA: Academic Press.

Labov, W. (1966). *The social stratification of English in New York City.* Washington, DC: Centre for Applied Linguistics.

Labov, W. (1969). Contraction, deletion, and inherent variability of English copula. *Language, 45*(4), 715-762.

Lorenz, D. (2013). Contractions of English semi-modals: The emancipating effect of frequency. NIHN Studies. Freiburg: Rombach.

Meillet, A. (1912). L'evolution des forms grammaticales. *Linguistique général et linguistique historique.* Paris: Champion. 130-148.

Pichler, H. (2009). The functional and social reality of discourse variants in a northern English dialect: I DON'T KNOW and I DON'T THINK compared. *Intercultural Pragmatics*, *6*(4), 561–596.

Sankoff, D., Tagliamonte, S., & Smith, E. (2012). *Goldvarb Lion: A variable rule application for Macintosh*. Department of Linguistics, University of Toronto.

Scheibman, J. (2000). *I dunno*: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics, 32,* 105-124.

Tagliamonte, S. & D'Arcy, A. (2007). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change, 19,* 199-217.

Thompson, S., & Mulac, A. (1991). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In E. C. Traugott and B. Heine (Eds.), *Approaches to Grammaticalization*, vol. 2, 313-330. Amsterdam: John Benjamins.

Torres Cacoullos, R., & Walker, J. (2009). The Present of the English Future: Grammatical Variation and Collocations in Discourse. *Language, 85*(2), 321-354.

Walker, J. (2010). *Variation in linguistic systems.* New York, NY: Routledge.

Weatherall, A. (2011). I don't know as a Prepositioned Epistemic Hedge. *Research on Language & Social Interaction*, *44*(4), 317-337.

# The role of storytelling in language learning: A literature review

**Claudio Rezende Lucarevschi**
University of Victoria
*clucarezende@gmail.com*

Storytelling is one of the oldest forms of human communication, and much has been said in the literature about its effectiveness as a pedagogical tool in the development of language skills in first (L1) language, and also in a foreign or second language (L2), regardless of learners' age or background (e.g. Isbell, Sobol, Lindauer & Lowrance, 2004; Cameron, 2001). Furthermore, storytelling is even claimed to be more effective in language teaching than traditional teaching materials, such as textbooks. Indeed, studies generally believe that effectiveness of storytelling relies on the fact that it is fun, engaging and highly memorable, raising learners' interest in listening to stories, as well as in speaking, writing and reading about them (e.g. Atta-Alla, 2012, Kim, 2010; Wajnryb, 2003). However, the studies in the literature generally lack specifics such as how exactly the effects of storytelling were measured or what specific skills are benefited from the use of storytelling, for example. Furthermore, the vast majority of the studies do not investigate any potential negative impact of the use of storytelling on language learning, giving us the impression that it is a pedagogical instrument that only brings a positive contribution to L2 learning. This literature review aims to provide an overview about what empirical studies say about the effects of storytelling on the development of language skills in L2, how storytelling compares to other teaching methods in its effectiveness, and 3) identify gaps in the current literature that should be addressed by future research. Addressing these questions will provide researchers and teachers with a clearer understanding about the role of storytelling in the language classroom, and, consequently, help them improve their teaching skills.
*Keywords: stories; storytelling; language skills; improvement; traditional teaching methods*

## Preamble

One month before I moved to Canada to start my PhD studies in Linguistics, in one of my language classes, in Brazil, I decided to tell my students a story for the first time, in order to see how they would respond to it. For this experience, I decided to choose the most challenging class I had, whose students had frequent behavioural problems and showed a lack of interest in participating in classroom activities and in learning languages, in order to investigate what the impact would be. I decided to tell my students a fictitious story, that presented values (i.e. trustworthiness and honesty) and emotions (i.e. love and hate). While I was telling my students the story, I noticed that all of them were attentively listening

to the story. In addition, after I finished telling the story, some of my students started to ask questions about it, and others demonstrated a strong interest in telling their own stories, based on similar experiences they had gone through. At that moment, I perceived that storytelling had a strong power to hold students' attention and encourage them to actively participate in oral and written activities related to the story they heard. I then started wondering whether the use of storytelling in the L2 classroom would also have a positive impact on helping students develop their language skills. This literature review has therefore offered me the opportunity to investigate the existing empirical evidence on the role and impact of storytelling in developing students' language skills in L2.

## 1 Introduction

Storytelling is one of the oldest forms of human communication, being used for entertainment, and for the promotion of education and cultural values. Furthermore, it is considered to be a very effective pedagogical instrument in the development of language skills in first (L1) language, and also in a foreign or second language (L2), regardless of learners' age or background (e.g. Isbell et al., 2004; Cameron, 2001). Indeed, storytelling is claimed to be more effective in language teaching than traditional language teaching methods, through the use of textbooks, as it is said to work on language skills in a fun, engaging and contextualized manner, consequently raising learners' interest in listening to stories, as well as in speaking, writing and reading about them (e.g. Atta-Alla, 2012, Kim, 2010). Moreover, storytelling is claimed to be very memorable to learners, helping them learn and retain vocabulary, grammatical structures and pronunciation (e.g. Wajnryb, 2003). But is there empirical evidence on such effectiveness of storytelling in the enhancement of language skills specifically in L2? What do studies say about the effectiveness of storytelling activities, compared to other language L2 teaching activities that do not use storytelling? Do studies report any negative impacts of the use of storytelling in the L2 classroom?

This literature review aims to investigate what empirical studies say about the effects of storytelling on the development of language skills in L2, and look for answers to questions such as why storytelling is used in the language classroom, how it is used, who tells stories to whom, and also whether studies refer to any negative impact storytelling may have on L2 learning. Addressing these questions will provide researchers and/or teachers with evidence on the role of storytelling in L2 learning, and enable them to improve and expand their pedagogical skills. The works to be reviewed here will focus on the effects of storytelling on L2 child and adult learners from different L1 backgrounds and of different age groups, for the purpose of verifying the impact of storytelling on these types of learners, who represent a large and increasing number of learners all over the world. This review also aims to compare the effects of different types of stories (i.e. stories about personal experiences, fictitious stories, folk tales) and different story formats (i.e. paper-based storytelling, and digital storytelling) on

L2 learning, in order to try to identify which ones have been most effective in the teaching of L2 and why. The aim will be to discuss the findings of studies which use different types of research designs (i.e. qualitative, quantitative, or mixed), for the purpose of examining the challenges posed, and the outcomes suggested by such designs. The argument presented is that while storytelling may have a positive impact on L2 learning, there may be aspects of storytelling and indeed potential negative impacts that have not yet been fully explored. Based on the existing evidence in the literature, this review will address the following questions:

1. Does storytelling improve L2 learners' receptive (listening and reading) and productive (writing and speaking) skills?
2. How does storytelling compare to other teaching methods in its effectiveness?
3. What kinds of gaps are there in the current literature that should be addressed by future research?

The review will start by providing some working definitions for storytelling and outlining the different types of stories and story formats that have been used in the L2 classroom (Section 2). The review will then analyze and discuss what qualitative and quantitative empirical studies say about the effects of storytelling on the development of L2 learning (Section 3), offer suggestions for future research (Section 4), and provide conclusions on the importance of storytelling (Section 5).

## 2 Storytelling: Definitions, types and formats

This section will discuss working definitions for storytelling, as well as the types of stories and story formats used in the L2 classroom, for the purpose of defining the terms and parameters to be used in this review of the literature.

### 2.1 Definitions

Researchers in the field of language teaching generally define storytelling in terms of how it works or what it does to promote communication between storytellers and story listeners. For example, Dyson and Genishi (1994) suggest that storytelling is a process where a teller uses a narrative structure, vocalization, and/or dramatic and mental imagery to communicate with an audience, who also uses mental imagery to provide the teller with verbal and non-verbal feedback. Likewise, Hsu (2010:7) defines storytelling as "the use of voice, facial expressions, gestures, eye contact, and interaction to connect a tale with listeners". A tale is created through the interaction between the storyteller and the audience. Thus, while the storyteller uses his/her voice and gesture to convey a story, the audience physically reacts to it by either squinting, staring or smiling, providing the storyteller with feedback on how storytelling is being received.

In terms of content or substance, storytelling is defined by McDrury and Alterio (2003) as:

> Uniquely a human experience that enables us to convey, through the language of words, aspects of ourselves and others, and the worlds, real or imagined, that we inhabit. Stories enable us to come to know these worlds and our place in them given that we are all, to some degree, constituted by stories (p.31).

An apparent contradiction seems to be found in the above definitions. Although Hsu's definition limits storytelling to an oral activity, the definition presented by McDrury and Alterio seems to offer room for other types of stories as well. This review will refer to storytelling as a creative human experience that allows us to refer to ourselves, to other people, to cultures in general, and also to real or imaginary worlds, through the language of words. In addition, this review will refer to storytelling as a process where a teller uses a narrative structure, vocalization, and/or dramatic and mental imagery to communicate with an audience, who also uses mental imagery to provide the teller with verbal and non-verbal feedback. This definition will be used in this review, because it includes form and content aspects of storytelling, and summarizes the key elements shared across the language literature.

Some researchers consider storytelling to be a different activity from reading aloud, whereas others do not make such distinction. Groeber (2007), Mello (2001) and Wang and Lee (2007), for instance, argue that, in storytelling, the teller focuses on the main message(s) of the story, and is free to use language improvisation, vocalization, mimetic action and his/her creativity to convey a story message to his/her audience. However, during the process of reading aloud a reader reads aloud the exact words in a given story or presents a memorized text to the audience. Nevertheless, most studies do not distinguish between these two forms of storytelling (e.g. Nicholas, Rossiter, & Abbott, 2011; Speaker, Taylor, & Kamen, 2004; Huang, 2006; Elkkiliç & Akça, 2008), and use storytelling as a broad term, which includes reading aloud. Like most studies in the literature, this review will consider reading aloud as part of a storytelling activity. Although in reading aloud a reader is expected to speak the exact words in a given story or present a memorized text to the audience, he/she can still use elements such as mimetic action and creativity to convey a story message to an audience, making the supposed differences between storytelling and reading aloud very slight. In order to verify whether there is a significant difference between storytelling and reading aloud, to justify a formal distinction between them, further research is needed to compare the effects of storytelling and reading aloud on the same group or similar groups of L2 learners. Such research would verify how learners respond to these two different storytelling approaches. In the absence of such evidence, the definition of storytelling used here will include reading aloud activities.

## 2.2 Story types and formats

Researchers use different types of stories and story formats to investigate the role of storytelling in the development of skills in L2. Differences in story types and formats play an important role in showing how L2 learners respond to the differences and, consequently, in helping teachers and/or researchers have a clear understanding on whether the story type or format used in the classroom matters and why. For example, with respect to story types, Huang (2006) studied the use of fairytales (stories that contain magic, fairies) to investigate the role of storytelling in English in the enhancement of the speaking skills of a group of young English as a Foreign Language (EFL) learners. In contrast, Cary (1998) used one fable (a short story that uses animals that talk or behave like humans as characters, to convey a moral), one fairytale, and two folktales (tales that are part of a culture or tradition and are orally told) to investigate the impact of storytelling on the development of the language skills of child learners of Spanish as a foreign language. Nicholas, Rossiter, & Abbott (2011) studied the impact of the use of stories of personal experiences on the development of learners' language skills in English as a Second Language. Likewise, Ko, Schallert, & Walters (2003) focus on the use of stories of personal experiences as storytelling tasks. That different studies have used different kinds of stories is a factor that must be taken into account when analyzing how learners respond to storytelling, and also in showing us the ways in which such different types of stories impact L2 learning and why.

With respect to story format, although studies widely agree that stories, in general, play a role in facilitating the improvement of language skills in L2, there is a lack of research comparing the effectiveness of printed and digital story formats on the development of L2 skills in a particular group of subjects. Indeed, even studies that use more than one type of story (e.g. Cary, 1998) do not analyze or discuss whether there was a particular type of story that was more effective and why. This would be important to show language teachers and/or researchers which type(s) of stories work best in a particular context, and with what kinds of learners (i.e. child learners, adult learners).

In addition to the more traditional printed stories and the use of oral stories (with no reference to printed texts), digital stories are also currently being used in the L2 classroom. Indeed, digital storytelling consists of computer-based programs that offer learners the opportunity to create their own stories through the use of spoken narratives, in which learners may even record stories using their own voices, visuals, soundtracks, and other types of technologies to share a story that is usually (but not necessarily) based on personal experiences or memories (Normann, 2010; Zheng, Yang, Zhaig, Wang, Sun & Xue, 2011). Furthermore, digital storytelling is claimed to be a motivating, efficient and interactional pedagogical tool, enabling learners to share their stories and provide one another with feedback on their content or performance in a fun and engaging manner (Porter, 2005; Sadik, 2008; Yang & Wu, 2012). Thus, the use of digital stories has also become a common procedure in studies examining the effects of

storytelling on the language learning process. Despite the fact that digital storytelling is considered to have a positive impact on language learning, through the engaging and technologically-advanced digital resources it offers (e.g. Skinner & Hagood, 2008; Tsou, Wang & Tzeng, 2006), traditional printed-format stories are claimed to be as effective as digital stories in the development of language skills. In other words, the value of storytelling in promoting the development of L2 learners' language skills is assumed to not be much influenced by the story format (i.e. printed and digital) used in the classroom.

The next section will analyze and discuss what questions qualitative and quantitative studies have investigated, and their findings, about the effects of storytelling on L2 learning. The aim of the following section will be to explore what kinds of empirical evidence on the feasibility and impact of different kinds of storytelling are available to teachers and researchers working with L2 learners. It will be followed in section 4 by a discussion of the gaps left by these studies, and potential areas for future research.

## 3 What do we know about the role of storytelling in the development of L2?

There is a considerable number of studies which investigate the effects of storytelling on the development of receptive and productive skills in L2 (e.g. Ajibade & Ndububa, 2008; Brown & Hirata, 2007; Cameron, 2001; Costenaro, 2008; Ellis, 2009; Hughes, 2009; Rachmawaty & Hermagustiana, 2010; Peck, 1989; Speaker et al., 2004; Sue & Bayley, 2005; Wajnryb, 2003; Wright, 2007), as well as on the development of grammar, vocabulary and pronunciation in L2 (e.g. Bardon-Harlig, 1995; Beaken, 2009; Madros, 2010; Mattheoudakis, Dvorakova and Láng, 2008; Wajnryb, 2003). Oral narratives are held to "divide text into tone units, identify tonic syllables on significant information, identify information that is already known and therefore non-prominent, identify points where the storyline changes direction - where the high fall (new information) or the fall-rise (familiar information, but with a new viewpoint) should be used" (Beaken, 2009, p.349). Thus, storytelling is claimed to help young learners become aware of rhythm and intonation features (Mattheoudakis, Dvorakova & Láng, 2008).

This section will review a range of qualitative and quantitative empirical studies, which may not be directly comparable as they use varying methods, participants of different age groups (i.e. child and adult learners) and from a variety of L1 backgrounds. Notwithstanding their differences, I will assess what kinds of insights these studies offer into the role of storytelling in the development of L2 language skills. I will start by reviewing studies which investigate the role of storytelling in the development of specific subsets of language skills, such as reading and speaking, and then I will examine studies which focus on the role of storytelling in developing the four language skills (reading, listening, reading and writing) as a whole.

### 3.1 Storytelling and the development of specific language skills

Some studies have focused on whether storytelling plays a role in developing specific sets of skills in L2, such as reading and speaking. To illustrate, Hsu (2010) performed a ten-week quantitative study on the role of English storytelling instruction in the development of speaking skills of a group of twenty-five grade five and twenty-five grade six elementary school students, who are native speakers of Taiwanese and study English as a foreign language. This small-scale study included one experimental group and one control group. The experimental group received instructions through the use of storytelling activities, as well as through access to an English textbook. The control group, however, only received instructions through the same English textbook. Both groups were pre-tested and post-tested with regards to their oral language complexity. No significant differences were found between the groups during the pre-test. However, in the post-test, the experimental group was shown to produce longer, more complex oral sentences in L2 than the control group. Hsu (2010) suggests that this happened because, during the study, learners did not only passively listen to stories, but also had the opportunity to retell the stories they heard, and practice the new vocabulary and sentence structures introduced to them. However, Hsu does not say what exactly he means by "more complex oral sentences" or what such sentences consist of. Therefore, it is not possible to know whether learners were assessed in terms of use of more advanced grammatical structures and/or vocabulary and whether their pronunciation and oral fluency, for instance, were also taken into account. Since Hsu does not focus on learners' oral grammatical accuracy or on their oral fluency, the term "more complex oral sentences" is vague and therefore limited in its use to language teachers and researchers.

Hsu's (2010) view on the importance of encouraging learners to be storytellers to promote the development of their oral skills in L2 is also shared by other researchers in the field. To illustrate, in another small-scale study, Afrilyasanti and Basthomi (2011) carried out qualitative research to investigate the role of digital storytelling in the enhancement of five grade eight EFL learners' oral production, focusing on their opinions regarding digital storytelling, as well as the quality of their digital storytelling production. The subjects attended seven meetings after school hours, although the length of each meeting is not specified. The researchers found that the learners were able to present their stories intelligibly and also through a smooth, spontaneous speech flow. By telling and retelling their stories, learners had many opportunities to practice their pronunciation. Furthermore, learners could assess their own fluency by listening to their recorded voices, although the study does not explain how exactly learners assessed their fluency and whether they received any descriptors or orientation from their teacher on how to self-assess their oral fluency level. Afrilyasanti and Basthomi also observed that the use of digital storytelling in the classroom motivated learners to voluntarily ask questions in class, respond to oral discussions, and introduce new topics during EFL classes. These authors

concluded that digital storytelling encourages learners to actively participate in speaking activities during their L2 learning process, helping them improve their vocabulary, pronunciation and fluency. The study by Afrilyasanti and Basthomi does not address the question of whether storytelling activities had any negative impact on the development of oral production of any of subjects or even on the instructor during the study. Such information would be important to detect whether there are any specific areas or issues that need further attention from L2 teachers.

Besides investigating the effects of storytelling on the development of speaking skills, researchers have also investigated the role of storytelling in helping L2 learners develop their reading skills. For example, Huang (2006) performed a significant quantitative study on the effects of Contextualized Storytelling Approach (CSA) on the development of reading comprehension of a group of seventy-two EFL learners in Taiwan. CSA uses different types of extralinguistic resources such as objects, body language, music, sound effects, and visuals, to mention a few, to facilitate comprehension of language learning. The study consisted of comparing the performance of three different groups of twenty-four grade six learners each (two experimental groups and one control group), in terms of reading comprehension and word recall. The first group (control) was exposed to text-only stories; the second group was exposed to illustrated written stories, and the third group first listened to stories and then was exposed to illustrated written texts of the stories they had listened to. The results of the study showed that the third group outperformed the other two groups in its ability to retell stories, suggesting that CSA helped learners improve their language comprehension. Nevertheless, it is not possible to know what exactly in CSA plays a primary role in promoting L2 learners' reading comprehension. Is it the use of prompts in CSA, the oral stories or a combination of both? This issue needs to be further investigated. Nonetheless, Huang's study suggests that the use of illustrations helps facilitate and positively impacts L2 learners' reading comprehension.

Similarly, Chang (2010) investigated the effectiveness of storytelling on the development of the reading skills and story recall of a group of fifty-two grade five native speakers of Taiwanese who studied English as a foreign language, by comparing two different teaching approaches that use storytelling: the teacher-mediated storytelling approach (TMSA) and the computer-mediated storytelling approach (CMSA). In the TMSA, the researcher told the participants a story, interacting with them by using flashcards to teach vocabulary. After telling the story, the researcher briefly reviewed it with the participants, before they were exposed to a questionnaire and tests. In the CMSA, the researcher pre- taught the key vocabulary that would be found in the story, and then orally presented the story to the participants. After the presentation, the participants read the story, and explored the interactive features available on the computer. They then answered a questionnaire that focused on learners' preferences and views of the efficiency of the TMSA and CMSA approaches, and did the reading comprehension and the story recall tests. Learners were assessed in two different moments: the first

experiment took place on December 15th and 18th 2009, and the second experiment was carried out on March 23rd and 26th 2010. Results from the analysis of the questionnaire showed that, although learners liked both approaches, they believed that the CMSA was more effective in facilitating the improvement of their language skills. In the first experiment, the TMSA group scored higher than the CMSA group in story comprehension, whereas in the second experiment, which took place around thirteen weeks later, the CMSA group outperformed the TMSA in reading comprehension and story recall. The second experiment used exactly the same procedures as the first experiment. Although such results suggest that increasing exposure to CMSA has a positive impact on the improvement of such performance, it is not clear whether the better performance of the CMSA group in the second experiment occurred due to learners' longer exposure to the CMSA approach, whether it was due to the fact that learners, in both groups, generally had a personal preference for the CMSA approach or even whether both exposure and preference to CMSA played a role.

As well as these studies that focus on the impact of storytelling on specific L2 language skills such as speaking, listening and reading, researchers have also investigated the relationship between storytelling and the development of language skills as a whole, as described in the following section.

### 3.2 Storytelling and the development of language skills as a whole

A range of empirical studies investigates the effects of storytelling on the development of receptive and productive language skills as a whole. For instance, Kim (2010) performed a six-week study to investigate the role of storytelling in the development of language skills of adult learners of English as a Second Language (ESL), and also the impact of learners' interest in learning ESL on the enhancement of their language skills. Five undergraduate and graduate learners, at basic English level, were used as participants in this small-scale study. The study was conducted twice a week (one hour per session), and data were collected and analyzed qualitatively (i.e. questionnaires, oral interviews and field notes), and also quantitatively. Learners were exposed to different types of stories, such as fables, folktales and stories on personal experiences, and were given the opportunity to not only to listen to stories, but also to read stories aloud, answer oral and written questions about the stories they heard, as well as write and orally tell their own stories to their classmates and instructor. Results of the study showed that storytelling was an effective pedagogical instrument in improving the language skills of learners who demonstrated interest and pleasure in participating in storytelling activities, whereas little progress was observed in learners with no or little interest in storytelling. Kim noticed that two out of the five participants in his study did not enjoy storytelling activities, and that these learners had a lower English proficiency level, when compared to the participants who demonstrated a strong interested in storytelling. Indeed, during the study, the participants with a lower English proficiency level showed a lot of stress in having to tell stories and frustration at not being able to satisfactorily cope with what they were being asked

to do in class. Despite the fact that the study suggests that the stress experienced was caused by the learners' lower English proficiency level, it is possible that stress may have also been caused by the way storytelling was conducted in class. In sum, Kim's (2010) study demonstrates that storytelling may play an important role in helping learners improve their language skills in L2 but also that it may not help learners much if storytelling is not delivered at an appropriate language level to learners, causing them stress, frustration and lack of interest in storytelling activities.

Likewise, Atta-Alla (2012) carried out a seven-week study to investigate the role of storytelling in enhancing the language skills of a group of forty male and female ESL adult learners, between thirty and fifty years of age, who had used English as a foreign language (EFL) for six to twelve years. During the study, the learners were told fifteen different types of stories, including folktales and picture book stories, which contained repeated patterns, such as rhymes, and the repetition of words and sentences. After listening to stories, learners participated in different oral and written communicative activities which aimed to offer them the opportunity to work on receptive and productive skills. Learners were encouraged to write their own stories and retell them to their peers and to their teacher, for instance. Post-test results showed that learners demonstrated higher test results in the four language skills (reading, writing, speaking and listening) than in the pre-test, suggesting that storytelling plays a role in enhancing such skills. Nevertheless, the study by Atta-Alla (2012) presents some methodological challenges that need clarification before further conclusions can be drawn. For instance, the study does not state how long learners were exposed to storytelling activities or the total length of the study. In addition, the claim on the importance of storytelling in improving language skills relies on the results of a three-hour test on the four language skills. In other words, although learners participated in different oral and written storytelling activities, their performance was only assessed through a post-test. In order to have a broader view of the importance of storytelling in enhancing the four language skills in L2, learners should also be assessed through other methods, such as continuous classroom activities, and have the results of their three-hour test compared to their performance in the classroom activities.

The studies by Kim (2010) and Atta-Alla (2012) offer empirical evidence on the importance of storytelling in promoting the development of receptive and productive language skills, by providing learners with the opportunity to actively work on such language skills in a fun and engaging manner. Furthermore, Kim's (2010) study also speaks to the potentially negative impact that storytelling may have on L2 learning. However, despite the suggested importance of storytelling, the negative impact of storytelling on L2 learning still needs to be further investigated and properly addressed in the literature. Indeed, the vast majority of the studies in this area provide us with a 'rosy' picture of the role of storytelling in developing L2 learning, giving us the impression that it is a pedagogical instrument that only brings a positive contribution to L2 learning. Some of these issues will be addressed in section 4, below.

Having looked at the language skills that storytelling is claimed to improve, questions still remain as to why and how storytelling improves such skills. Empirical studies claim that storytelling improves language skills for three main reasons: because the use of stories is highly motivating; because they provide learners with comprehensible input; and because they promote social interactions. The following sections will evaluate these claims in turn.

### 3.3 Storytelling as a motivating tool

Storytelling is considered to be an effective instrument to boost learners' motivation to learn languages (e.g. Wajnryb, 2003; Wright, 1995), increasing their active participation in the language learning process. According to Wajnryb (2003), the content and meaning of stories are powerful enough to engage learners in the learning process, helping them improve their skills in L2. Similarly, Wright (1995) claims that storytelling increases learners' interest in listening and reading stories, playing an important role in improving their listening and reading skills in L2. Storytelling is, therefore, considered to be a powerful instrument to motivate learners to learn languages by engaging and encouraging them to actively participate in the language learning process.

This issue was examined in one of the most recent studies in the field. Yang and Wu (2012) carried out a significant, one-year quasi-experimental study of the effects of digital storytelling (DST) on the motivation of one hundred and ten grade ten EFL learners in Taiwan and on the improvement of their language skills in EFL. The participants were divided into an experimental group and a comparison group. Although both groups were pre and post-tested, and also exposed to the same language contents, assessment tasks and instructor, they used different teaching strategies. The experimental group was exposed to DST activities, through which learners were encouraged to actively participate in the language learning process, using creative thinking in the production of oral and written language. In contrast, the control group was mostly exposed to textbook readings and the listening of lectures on PowerPoint slides. Data were collected and analyzed qualitatively and quantitatively, through questionnaires, tests and descriptive statistical analysis.

Results indicated that the participants in the experimental group outperformed the control group in all three groups of variables measured: learning motivation, creative thinking and language skills and features (listening, reading, writing, grammar and vocabulary). Indeed, Yang and Wu suggest that DST played a very important role in improving learners' motivation and language skills by providing them with opportunities to work on language skills actively and creatively, through the use of modern and engaging digital resources. Moreover, learners exposed to DST were encouraged to use language in meaningful contexts, related to their personal experiences. Yang and Wu's study has made a substantial contribution to the literature, as it investigated the effects of DST on the motivation of a large number of L2 learners (n= 110), for a long period of time

(1 year), and by using mixed (qualitative and quantitative) research design.

In a smaller scale study, Elkkiliç and Akça (2008) conducted research on the role of storytelling in motivating a group of Turkish child learners to learn EFL. Twenty-one grade four elementary learners participated in this study. They were asked to read a series of statements in a questionnaire, and choose the statements which best reflected their opinion. In addition, they were asked open-ended questions regarding their reason(s) for studying EFL (in order to determine whether they were being intrinsically or extrinsically motivated to learn English), the activities they like doing in their classes, as well as the activities that they do not enjoy doing. According to Schunk (2008), extrinsic motivation occurs when a learner is motivated to do something for an external reward, such as higher grades or a job promotion, whereas intrinsic motivation refers to one's desire to participate in an activity, solely for the pleasure in doing the activity itself.

In Elkkiliç and Akça' study, most participants reported to enjoy the storytelling activities in their EFL classes, and learn something new from the stories they had been exposed to. Results indicate that learners perceive storytelling as an important motivational tool in EFL learning. Although learners' opinions on the importance of storytelling in motivating and helping them improve their language skills in L2 is important to teachers and researchers to show what learners' views are, it may not reflect how learners actually react or respond to storytelling activities in the classroom. Thus, in order to confirm whether learners' opinions are in agreement with the actual way they respond to storytelling activities, the kind of study carried out by Elkkiliç and Akça could also provide learners with storytelling activities and evaluate their motivation level during such activities. Such evaluation could be carried out through classroom observations, for example. We could then compare the findings of the questionnaire on learners' views to the classroom observations. This comparison would provide language teachers and researchers with more reliable, comprehensive conclusions on the role of storytelling in enhancing learners' motivation and language learning.

Taken together, these two studies seem to support the claims on the effectiveness of storytelling in motivating learners to learn L2 and in enhancing their receptive (i.e. reading and listening) and productive (i.e. speaking and writing) skills, by providing learners with the opportunity work on the referred language skills in fun, meaningful and engaging activities. Besides helping L2 learners develop their language skills as a motivating tool, storytelling is also claimed to help learners improve their language skills through social interactions, as assessed in the next section.

### 3.4 Storytelling and the development of language skills through social interactions

Researchers generally believe that storytelling plays an important role in the development of language skills in L2, by promoting social interactions and mutual

collaboration in the L2 classroom (e.g. Wajnryb, 2003). For example, storytelling is claimed to encourage learners to interact with each other by listening to and telling stories to each other and by reading each other's stories, in addition to encouraging them to interact with their teachers. Such interactions, which offer learners the opportunity to receive support from their teachers and also from classmates who may have a slightly higher knowledge of the L2 being studied, are also considered to help learners work within their current and potential level of development, or their Zone of Proximal Development (ZPD) (e.g. Ko, Schallert & Walters, 2003; Ryokai, Vaucelle & Cassell, 2003). As regards ZPD, it is a concept created by Vygotsky, and defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p.86). Thus, by interacting with teachers and/or more skilled peers, learners are believed to develop their language skills more satisfactorily than they would have done had they worked alone.

In terms of empirical studies, Li and Seedhouse (2010) carried out research to investigate the role of storytelling in the development of oral interactions in elementary learners of English as a Foreign Language (EFL) in Taiwan. The study compared learners' interactions in standard EFL classes (with the support of a textbook) to story-based classes. Two teachers from two different elementary schools participated in this two-month study. Each of them worked with a different class (aged 10) of 30 to 35 learners each. Twenty-six lessons were recorded and transcribed. Li and Seedhouse found an increase in oral production and expression of different language functions with the use of story-based lessons as compared with standard lessons, which could be due to an increase in interest and motivation to participate in story-based lessons, as evidenced by an increase in the learners' spontaneous oral interactions during these classes. Li and Seedhouse also reported that story-based lessons were more efficient in expanding learners' vocabulary when compared to standard lessons. The study by Li and Seedhouse seem to suggest that storytelling is an effective pedagogical tool in promoting social interactions in the L2 classroom, and also in developing different language areas, such as oral production and vocabulary.

Similarly, in a very short study, Ryokai et al. (2003) investigated whether interactions with a slightly more skilled partner in the target L2 during a storytelling activity play a role in helping twenty-eight 5-year-old learners enhance their language skills. The study used a female virtual peer (Sam), who looked like a 6-year-old child. This virtual peer was projected on a screen behind a castle, and interacted with the learners, by telling them stories and encouraging them to tell their stories as well. The virtual peer gradually used more advanced linguistic expressions during the storytelling, and her responses were controlled by a researcher who stayed behind the screen. The session lasted about fifteen minutes. Ryokai et al. (2003) noticed that the participants who played with Sam told stories using language structures and vocabulary similar to those used by the virtual peer. Although the results from this study suggest that interactions with

slightly more skilled interlocutors enhance learners' language skills, further research needs to be conducted before conclusions are drawn, as this study was conducted for only fifteen minutes. Therefore the study by Ryokai et al. does not provide any follow-up on whether the claimed benefit of using a virtual partner in storytelling activities to promote the development of language skills in L2 would sustain for a longer period. Moreover, in a fifteen-minute study it is hard to know whether the development of learners' skills occurred because they were exposed to a nice and different kind of peer, whether it was because of the use of storytelling, or whether learners were benefited by both (i.e. the quality and kind of interactions with the virtual peer, and the use of storytelling).

Results of the studies by Li and Seedhouse and Ryokai et al. suggest that storytelling helps learners improve their language skills, by increasing their social interactions in the classroom. In addition, some researchers have claimed that storytelling enhances L2 learners' language skills by providing them with comprehensible input, as discussed in the following section.

### 3.5 Storytelling as a tool to provide comprehensible input

Researchers generally agree that storytelling creates ideal language learning conditions, because it provides learners with *comprehensible input* (e.g. Brewster, Ellis & Girard, 2002, Hendrickson, 1992; Wajnryb, 2003). Wajnryb, for instance, claims that "the text of the story provides the potential for comprehensible input, that is language that is within the range of access of the learner" (p.7). Some empirical studies have investigated the role of storytelling in providing learners with comprehensible input and facilitating the development of skills in L2.

For example, Cary (1998) performed a qualitative study on the effectiveness of a Contextualized Storytelling Approach (CSA) in facilitating the comprehension of English, which was the participants' L2, as well as in impacting the quantity of L2 speaking in a group of twelve elementary Latino learners, whose L1 was Spanish. Three English-Spanish bilingual classrooms were analyzed during a period of four weeks. Data were collected through observations, field notes, recorded dialogues between the researcher and the three participant teachers, and through recorded informal dialogues between the teachers and the learners. Four traditional stories were selected for the study. Props, costumes, music, movement, and sound effects, as well as synonyms and paraphrasing, role-playing, and teacher-facilitated post-story discussion were used to promote comprehension of the oral stories. The results obtained showed that CSA increased learners' comprehension of L2. Indeed, Cary observed that learners were engaged to attentively listen to stories, which facilitated their comprehension. Furthermore, the quantity of L2 speaking was also increased after exposure to CSA stories. Despite the fact that the findings from Cary's research are limited by the lack of a control group, the study makes an important contribution to the literature, by showing that CSA facilitates learners' comprehension and also develops their listening and speaking skills in L2. The use of a control group would play a

relevant role in Cary's study, since it investigates the effects of a pedagogical instrument (CSA) on learners' comprehension and speaking skills in L2.

A further study on the role of storytelling in providing learners with comprehensible input was performed by Isbell et al. (2004) compared the effects of storytelling and reading aloud activities on the improvement of story comprehension of thirty-eight 3 to 5 year-old children during a 12-week study. The study used 2 groups: the first group was exposed to storytelling activities, whereas the second was exposed to reading-aloud activities. As mentioned in section 2, in storytelling the teller improvises what he/she will orally say, while in a reading aloud activity the story reader reads aloud the exact words of a given story or presents a memorized text to his/her audience. Results showed that although the storytelling group outperformed the reading aloud group in story retelling, and the reading aloud group outperformed the storytelling group in the development of a wordless picture book story, both storytelling and reading aloud played an important role in providing learners with comprehensible input. Indeed, storytelling and reading aloud are fun, engaging and meaningful tasks, providing learners with a clear understanding of the meanings they convey.

Clearly there is a large range of literature in the field that addresses the impact of storytelling on L2 learners. Despite the methodological (i.e. research design, type of story and story format) differences found across the reviewed empirical studies in section 3, all of the reviewed studies suggest that storytelling is a very important pedagogical tool in facilitating the development of specific sets of language skills in L2, such as speaking, listening or reading, as well as receptive and productive skills as a whole. The next section summarizes the conclusions that can be drawn from this range of reviewed studies.

### 3.6 The effectiveness of storytelling: Some conclusions

The studies discussed here show differences and similarities in the way they investigate the role of storytelling in developing L2 language skills. For example, the great majority of the empirical studies use children as subjects and, therefore, it may give us the impression that storytelling is most effective when used with children than in adults. Nevertheless, storytelling is also reported to be an effective pedagogical instrument to help adult learners enhance the learning of an L2 (e.g. Atta-Alla, 2012; Kim, 2010). Although studies explore things in different ways, some common threads can be seen across their findings, such as the importance of encouraging learners to actively participate in the language learning process by not only to listening to stories but also by discussing them and telling their own stories in class. Indeed, learners' active participation in storytelling activities is held to help learners develop their language skills in L2, since it promotes motivation to learn a second/foreign language and social interaction in the L2 classroom.

Despite their wide range of approaches, and the use of different types of stories (e.g. fables, fairy tales, stories about personal experiences) and story

formats (paper-based, oral and digital), few studies investigate the impact of different types of stories and story formats within the same group of learners, to verify how they compare in their effectiveness to improve language skills in L2. This is one example of an important gap in the literature, and one that needs further clarification. However, there are other areas that also require further investigation. The next section offers some suggestions for future research, based on the gaps found in this review.

## 4 Suggestions for future research

In undertaking this review, some gaps have been identified in the literature on the impact of storytelling on L2 learning. Such gaps fall into six main themes: 1) L2 learners' age groups; 2) learners' L1 background; 3) the negative impact storytelling may have on the improvement of L2 skills; 4) the role of storytelling in the development of oral grammatical accuracy in L2; 5) lack of specifics on the effects of storytelling on the development of pronunciation in L2; 6) the impact of storytelling on the curriculum and also on L2 teachers.

First, the vast majority of studies on the role of storytelling in the development of language learning investigate its effects on children (e.g. Afrylyasanti & Basthomi, 2011; Li & Seedhouse, 2010; Speaker et al. 2004; Tsou et al., 2006), a more rounded picture of the effects of story telling on L2 skills would take into account a much broader age rage. Consequently, there should be further research on the effects of storytelling on the enhancement of teenage and adult learners' foreign/second language skills.

Secondly, most of the reviewed studies use monolingual groups of language learners (learners with the same L1) and data were collected in L2 learners' countries of origin. This allows us to conclude that such studies give us a limited view of the role of storytelling in the development of language learning. In order to have a more comprehensive evaluation of its role in language learning, it would be relevant to investigate the effects of storytelling on multicultural language groups and in their target language countries. For example, it would be important to investigate the relationship between storytelling and the improvement of oral grammatical accuracy using adult multicultural groups.

Thirdly, with respect to story types and formats, since the great majority of the studies in this area only show storytelling as positively impacting on the development of language skills in L2, it would be relevant to investigate whether and how storytelling might negatively impact on L2 learning. For instance, could the theme of a story or how the story is told negatively affect L2 learning? If so, in what ways and why? Would two groups of L2 learners present any difference in the development of their skills in L2 if each of these groups were exposed to a different story format, type of story and/or story contents? In fact, no studies compare the effectiveness of the different types of stories and story formats used in storytelling activities on the development of L2 skills. For example, none of the reviewed studies investigates whether there is a particular type of story (e.g.

folktale, fairy tale, personal experience) or story format (e.g. oral, paper-based or digital) that impacts more positively on the development of L2 skills than another. In fact, the studies solely suggest that the referred story types and formats are all effective in helping L2 learners improve their language skills.

Fourthly, with respect to language skills, although studies have suggested that storytelling enhances learners' grammar abilities (e.g. Bardovi-Harlig, 1995; Hsu, 2010) while others report that storytelling improves speaking skills (Cary, 1998; Li & Seedhouse, 2010), it seems that there are currently no studies that have evaluated the relationship between storytelling and the development of oral grammatical accuracy. This is certainly a gap that requires further investigation, as grammatical accuracy is considered to play an important role in oral communication (Lee, C., 2009; Spratt, Pulverness, & Williams, 2011), and therefore future studies are warranted in order to study this relationship. Furthermore, evaluation of the effects of different types of narratives on the development of oral grammatical accuracy is also warranted. In particular, it would be important to investigate whether monolingual and multicultural groups of language learners perform differently depending on the type of narrative employed. Research in the area would be very informative to language teachers, mainly in the selection of the optimal contents and formats of stories to be used in storytelling activities during their L2 classes. Moreover, it would be relevant to conduct comparative studies on the effects of different types of narratives on the development of learners' receptive and productive skills. This type of research would be particularly important in helping teachers in the selection of the types of narratives they can use in their language classrooms.

Furthermore, the vast majority of the studies that refer to the effect of storytelling on the improvement of oral skills, claiming that storytelling improves pronunciation in L2, do not specify what exactly they mean by '*improve* pronunciation', or what they mean by 'pronunciation'. For example, does 'improve' mean that learners develop a more 'native-like' pronunciation? Does storytelling develop pronunciation at segmental level (i.e. the pronunciation of segments in a word), at suprasegmental level (i.e. word syllable stress, sentence intonation) or at both? Such specifics would be important to show language teachers, for example, in which pronunciation areas the use of storytelling is effective. Finally, based on the gaps found in the reviewed studies, the suggested future research will make an important contribution to clarifying important aspects, regarding the role of storytelling as a pedagogical instrument to help L2 learners effectively improve their language skills in L2.

Lastly, and perhaps most importantly for L2 teachers, given that language teachers are supposed to teach a variety of language topics and usually have little time to cover them all, how impactful is storytelling on the L2 curriculum and those who teach it? To illustrate, how much time do L2 teachers need to spend on learning storytelling techniques and getting prepared to adequately use storytelling in the L2 classroom? How much preparation time is required from teachers to prepare storytelling activities and implement them in the L2 classroom? How much investment in teacher training programs would be

necessary to enable teachers to appropriately use and explore storytelling techniques in the classroom? These questions address specific and relevant issues of crucial interest to L2 teachers who may wish to incorporate storytelling into their curriculum.

## 5 Conclusions

This literature review provided an overview of the importance of storytelling in language learning by presenting and discussing several qualitative and quantitative studies that investigated the role of storytelling during language learning, particularly in L2 learning. Overall it can be concluded on the basis of these studies that storytelling is an effective pedagogical instrument to enhance learners' language skills in L2. For example, storytelling is suggested to help L2 learners enhance their language skills, because it offers them opportunities to work on specific sets of language skills (i.e. reading, listening and speaking) or work on receptive and productive skills as a whole, through meaningful and contextualized activities. Furthermore, the reviewed studies suggest that encouraging learners to be storytellers plays a very relevant role in promoting language learning, by encouraging them to actively participate in the learning process. Studies also suggest that storytelling promotes the development of language skills in L2, because it boosts learners' motivation to learn a second or foreign language, promotes social interactions among learners and between learners and teachers in the L2 classroom, helping learners keep within their ZPD, and also provides them with comprehensible input. Indeed, the reviewed studies suggest that storytelling plays an important role in facilitating comprehension because stories are highly contextualized, through the use of illustrations, verbal repetitions and/or body language, for instance, and also because they are fun and engaging.

Nevertheless, as discussed in the previous section, the role of storytelling in developing L2 language learning needs to be further investigated, as the literature lacks studies on whether storytelling may negatively impact on language learning. For example, studies do not explore whether the type of story, the story format, or the way storytelling activities are developed in class affect learners' performance in L2 and the development of their language skills. Moreover, studies do not assess whether storytelling has any negative impact on language teachers and on the L2 curriculum by demanding a considerable amount of preparation time and investment on the training of teachers to appropriately use and explore storytelling techniques in the classroom. Research exploring potential drawbacks in the use of storytelling would certainly benefit language teachers and the literature in general, by offering a comprehensive analysis and evaluation of the role of storytelling in the development of L2 language skills.

## References


Afrilyasanti, R., & Basthomi, Y. (2011). Digital Storytelling: A case study on the teaching of speaking to Indonesian EFL students. Language in India: *Strength for Today and Bright Hope for Tomorrow*, *11*, 81-91.

Ajibade, Y., & Ndububa, K. (2008). Effects of word games, culturally relevant songs, and stories on students' motivation in a Nigerian English language class. *TESL Canada Journal/Revue TESL Du Canada 27, 25*(2), 27-48.

Atta-Alla, M. (2012). Integrating language skills through storytelling. *English Language Teaching (5)*, 12, 1-13.

Bardovi-Harlig, K. (1995). A narrative perspective on the development of the tense/aspect system in second language acquisition. *SSLA, Cambridge University Press, 17*, 263-291.

Beaken, M. (2009). Teaching discourse intonation with narrative. *ELT Journal. Oxford University Press*, *63*(4), 342-352.

Brewster, J., Ellis, G., & Girard, D. (2002) *The Primary English Teacher's Guide*. New Edition. England: Pearson Education.

Brown, T., & Hirata, E. (2007). Are you sitting comfortably? The role of storybooks in primary English education. *Bulletin of Faculty of Education, Nagasaki University: Curriculum and Teaching, 47,* 117-127.

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.

Cary, S. (1998). *The effectiveness of a contextualized storytelling approach for second language acquisition*. (Unpublished doctoral dissertation). University of San Francisco, California.

Chang, J. (2010). *The comparison between the teacher-mediated storytelling approach and the computer-mediated storytelling approach on EFL children's English reading comprehension and story recall in Taiwan*. (Unpublished master's thesis). National Pingtung University of Education, Taiwan.

Costenaro, V. (2008). Language acquisitional storytelling: psycholinguistics in action in the Italian EFL classroom. *Studi di Glottodidattica*, *3*, 65-88.

Dyson, A., & Genishi, C. (1994). Introduction: The need for story. In Dyson, A., & Genishi, C. (Editors). *The need for story: Cultural diversity in classroom and community*. Illinois: The National Council of Teachers of English.

Elkkiliç, G., & Akça, C. (2008). Attitudes of the students studying at Kafkas university EFLclassrooms towards storytelling and motivation. *Journal of Language and Linguistic Studies*, *4*(1), 1-27.

Ellis, B. (2009). *Content area reading, writing, and storytelling: a dynamic tool for improving reading and writing across the curriculum through oral language development*. Connecticut: Teacher Ideas Press.

Groeber, J. (2007). *Designing and using rubrics for reading and language arts, K-6*. Second edition. California: Corwin Press.

Hendrickson, J. (1992). Storytelling for foreign language and linguistics. Washington, D. C.: *ERIC Clearing House on Language and Linguistics. (Eric Document Reproduction No. ED 355-824)*, 1-25.

Hsu, Y. (2010). *The influence of English Storytelling on the Oral language complexity of EFL primary students*. (Unpublished master's thesis). National Yunlin University of Science & Technology, Yunlin.

Huang, H. (2006). The effects of storytelling on EFL young learners' reading comprehension and word recall. *English Teaching & Learning*, *30*(3), 51-74.

Hughes, K. (2009). How social stories support early literacy learning of toddlers. In Meier, D. (Ed.). *Here's the story: Using narrative to promote young children's language and literacy learning*. New York: Teachers College Press.

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, *32*(3), 157-163.

Kim, M. (2010). The effects of storytelling on adult English language learners. *Linguistic Research, 27*(3), 447-473.

Ko, J., Schallert, D., & Walters, K. (2003). Rethinking Scaffolding: Examining Negotiation of Meaning in an ESL Storytelling Task. *TESOL Quarterly*, *37*(2), 303-324.

Lee, C. (2009). *Thoughts on foreign language education: Introduction to babble training method*. Bloomington: Author House.

Li, C., & Seedhouse, P. (2010). Classroom interaction in story-based lessons with young learners. In Robertson, P.; Nunn, R. (Eds.). *The Asian EFL Journal Quarterly*, *12*(2) 288-312.

Madros, J. (2010). *Traditional Oral Storytelling in the Classroom*. (Unpublished master's thesis). University of Alaska, Fairbanks.

Mattheoudakis M., Dvorakova, K., & Láng, K. (2008). Story-based language teaching: an experimental study on the implementation of a module in three European countries. In Nikolov, M., Djigunovic, J., Mattheoudakis, M., Lundberg, G., & Flanagan, T. (Eds.). *Teaching Modern Languages to Young Learners: Teachers, Curricula and Materials*. Council of Europe, 59-76.

McDrury, J., & Alterio, M. (2003). *Learning through storytelling in higher education.* Sterling, VA: Kogan Page Limited.

Mello, R. (2001). Storytelling and literacy learning: A review of the literature. *Wisconsin English Journal 43*(2), 20-28.

Mustaba, M. (2008). Storytelling: A boon to children's language development. *Journal of Urban Education: Focus on Enrichment, 5* (1) 68-73.

Nicholas, B., Rossiter M., & Abbott, M. (2011). The Power of story in the ESL classroom. The Canadian Modern Language Review/La Revue canadienne des langues vivantes, *67*(2), 247-268.

Normann, A. (2011). *Digital storytelling in second language learning: A qualitative study on students' reflections on potentials for learning*.

(Unpublished master's thesis). Norwegian University of Science and Technology, Trondheim.

Peck, J. (1989). Using storytelling to promote language and literacy development. *The Reading Teacher*, *43*(2), 138-141.

Porter, B. (2005). *Digitales: The art of telling digital stories*. Denver: Bernajean Porter Consulting.

Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning*, *19*, 195-208.

Sadik, A. (2008). Digital storytelling: a meaningful technology-integrated approach for engaged student learning. *Educational Technology Research and Development, 56*(4),487–506.

Schunk, D. (2008). *Learning theories: An educational perspective*. 5th Edition. New York: Prentice Hall.

Skinner, E., & Hagood, M. (2008). Developing literate identities with English language learners through digital storytelling.*The Reading Matrix, 8(2)*, 12-38.

Speaker, K., Taylor, D., & Kamen, R. (2004). Storytelling: Enhancing language acquisition in young children. *Education (Chula Vista), 125*(1), 3-14.

Spratt, M., Pulverness, A., & Williams, M. (2011). *The TKT teaching knowledge test course: Modules 1, 2 and 3*. Cambridge: Cambridge University Press. Sue, P., & Bayley, R. (2005). *Early literacy fundamentals: a balanced approach to language, listening, and literacy skills*. Ontario: Pembroke Publishers.

Tsou, W., Wang, W., Tzeng, Y. (2006). Applying a multimedia storytelling website in foreign language learning. *Computers & Education*, *47*, *Elsevier*, 17–28.

Vygotsky, L. (1978). *Mind in society: the development of higher psychological processes*. Cambridge: Harvard University Press.

Wajnryb, R. (2003). *Stories: Narrative activities in the language classroom*. Cambridge Handbooks for Language Teachers. Cambridge: Cambridge University press.

Wang, F., & Lee, S. (2007). Storytelling is the bridge. *International Journal of Foreign Language Teaching*, *3(*2), 30-35.

Wright, A. (1995). *Storytelling with children*. Oxford: Oxford University Press.

Wright, A. (2007). Stories and who we think we are. In Rubio, F. (Ed.). *Self-esteem and foreign language Learning*. Cambridge: Cambridge University Press.

Yang, Y., & Wu, W. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & Education 59*, 339–352.

Zheng, J., Yang, J., Zhang, H., Wang, L., Sun, Z., & Xue, H. (2011). Application of digital storytelling in comprehensive practice activity curriculums. In Liu, C., Chang, J., & Yang, A. (Eds.). *Communications in computer and information science. Information, computing and applications. Second*

*International Conference, ICICA 2011, Qinhuangdao, China, October 2011. Proceedings, Part I*. New York: Springer. 406-414.

# The /s/-/ʃ/ confusion by Japanese ESL learners in grapheme-phoneme correspondence: bias towards [s] and <s>

**Akitsugu Nogita**
University of Victoria
*akitsugu@uvic.ca*

It is generally believed that Japanese English-as-a-second-language (ESL) learners tend to pronounce English /siʲ, sɪ/ as [ʃiʲ, ʃɪ], such as *see* and *sip* as *she* and *ship* respectively, and these errors are typically attributed to the Japanese phonotactic constraint *[si(:)]. However, Nogita (2010) reveals that such errors are due to their misinterpretation of the spellings of <s> and <sh>, not due to articulatory and perceptual difficulties. In this present study, I further reinforced Nogita's (2010) argument by conducting a reading task in which 42 Japanese ESL learners read nonsense words containing the graphemes <s> and <sh>, and a spelling task in which they spelled nonsense words containing the sounds [s] and [ʃ]. In the reading task, I found Japanese ESL learners' strong tendency of mispronouncing the grapheme <sh> as [s], presumably because they assumed that [s] sounded more English-like. In the spelling task, they misspelled the sound [ʃ] as <s> more frequently than [s] as <sh>, presumably due to *kunrei-shiki* Japanese romanization interference. Moreover, 29 participants' grapheme-to-phoneme and phoneme-to-grapheme conversion patterns were not consistent, indicating that they had not acquired the English grapheme-phoneme correspondences, <s>-/s/ and <sh>-/ʃ/.
*Keywords: /s/-/ʃ/ confusion; <s>-<sh> confusion; second language grapheme-phoneme correspondence*

## 1 Introduction

### 1.1 Background

It is generally believed that Japanese English-as-a-second-language (ESL) learners tend to pronounce English /siʲ, sɪ/ as [ʃiʲ, ʃɪ][1], such as *see* and *sip* mispronounced as *she* and *ship* respectively. Such errors are typically attributed to the Japanese phonotactic constraint in which [ʃ] is an allophone of /s/ before /i(:)/ (e.g., Avery & Ehrlich, 1992). However, some phonologists stated that in Japanese, [si(:)] and [ʃi(:)] are marginally phonologically contrastive (e.g. Vance,

[1] Slashes / / are used for phonemes or mental representations. Square brackets [ ] are used for phonetic realizations. In some cases these brackets can be interchangeable.

2008). Likewise, according to Matsuzaki (1993), not all linguists agree that Japanese [si(:)] and [ʃi(:)] are non-contrastive. My previous study, Nogita (2010), also proves that this phonological contrast does exist in Japanese at least in peripheral vocabulary, as in <水橋パルスィ>[2]-/midzuhaʃi parusi(:)/ (a name of a game character), as well as in a near minimal pair <らしい> /raʃii/ 'seem' and <ばらスィー> /barasi:/ 'Barasī (a pseudonym of a Japanese manga artist). Moreover, in Nogita (2010), I collected data from 93 monolingual standard Japanese speakers (aging from 17 to 89) to demonstrate that all the participants can distinguish [si] and [ʃi] in both production and perception if these sounds are in Japanese contexts, suggesting that Japanese ESL learners' /s/-/ʃ/ confusion in English cannot be an articulation or perception issue. In addition, Nogita (2010) also demonstrates that Japanese ESL learners' /s/-/ʃ/ confusion in English contexts can be easily corrected when learners are only taught the Grapheme-Phoneme Correspondence (GPC) rules (i.e. <see>-/siʲ/, <she>-/ʃiʲ/, <si>-/sɪ/, <shi>-/ʃɪ/ and so on) without any articulation training. Indeed, their /s/-/ʃ/ confusion may partially be a phonological issue, that is, since the functional load of the Japanese /si(:)/-/ʃi(:)/ contrast is very low, Japanese ESL learners may not pay much attention to this contrast in English. However, I assume that their English spelling knowledge plays a much bigger role. While my previous study, Nogita (2010), reveals that a training of spelling knowledge dramatically reduces their /s/-/ʃ/ production errors, I have not discovered how Japanese ESL learners read the spellings <s> and <sh> as well as how they spell /s/ and /ʃ/ when they encounter unfamiliar words. Therefore, this present study aims to fill in these gaps in order to further support Nogita's (2010) argument that Japanese ESL learners' /s/-/ʃ/ confusion is more of an orthographic issue, rather than a phonological issue.

In L2 learning, what must be avoided is that Japanese ESL learners' English /s/-/ʃ/ confusion is misdiagnosed as a phonological issue and an unnecessary pronunciation training is provided only to further confuse learners, just as misdiagnosis of illness and medication errors only make patients suffer from side effects. To avoid such misdiagnosis, there are good reasons to examine to what extent Japanese ESL learners understand L2 GPC rules.

### 1.2 Japanese romanization regarding /s/ and /ʃ/ and phonetics

If this is in fact an orthographic issue, learners' L1 spelling, more specifically *rōmaji* (Japanese romanization) spelling, may interfere in L2, so I will briefly discuss Japanese romanization. There are two major types of *rōmaji* regarding [ʃi]: [ʃi] is spelled as <si> in *cabinet-ordered rōmaji* (or *kunrei-shiki rōmaji,* implemented in 1937) while it is also spelled as <shi> in *Hepburn rōmaji* (or *Hebon-shiki rōmaji* adopted in 1908 by an American missionary James C. Hepburn) (Taylor & Taylor, 1995). A difference between these two types is seen in some coronal obstruents, which reflect inconsistencies in some spellings in the

[2] Angle brackets < > are used for graphemes or written forms.

Japanese *kana* script (one of the scripts in the Japanese writing system). In the *kana* chart (or the 50-sound chart), the moras /sa, ʃi, sɯ, se, so/, for example, are aligned in the same consonant column despite occurrence of two different consonants, /s, ʃ/. This inconsistency mirrors historical sound changes[3]. In *cabinet-ordered rōmaji*, /sa, ʃi, sɯ, se, so/ are spelled as <sa, si, su, se, so> to be consistent with the *kana* chart, whereas in *Hepburn rōmaji*, these moras are spelled as <sa, shi, su, se, so> to be consistent with sounds (Taylor & Taylor, 1995).

Typing the Japanese *kana* script with a computer is typically based on both types of Japanese romanization. For instance, pressing either "s-i" keys (*cabinet-ordered rōmaji*) or "s-h-i" keys (*Hepburn rōmaji*) makes the *kana* letter <し> corresponding to [ʃi]. This suggests that native Japanese speakers are likely to be familiar with both spellings of [ʃi]. As for the rōmaji spelling for [si], according to the Ministry of Education, Culture, Sports, Science and Technology-Japan (2009), there is no official rule, but according to the Department of English Language in The University of Tokyo (2009), the recommended rōmaji spelling for [si] is <si>. This means that the rōmaji spelling <si> corresponds to both [si] and [ʃi], and at the same time, the sound [ʃi] corresponds to both <si> and <shi>, as it is shown in Figure 1 below[4]. This complicated relationship might influence their L2.

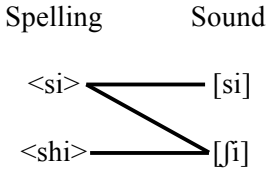



*Figure 1. Japanese rōmaji spelling-sound correspondence*

Incidentally, phonetic qualities of English /s/ and /ʃ/ and their Japanese counterparts are not the same. According to Pan, Utsugi, and Yamazaki (2004), compared to the English /ʃ/, the Japanese /ʃ/ is articulated further back and the front part of the tongue is higher. As well, the Japanese /ʃ/ does not have lip rounding and dorsum elevation, unlike the [ʃ] English counterpart (Pan et al, 2004). Therefore, the Japanese /ʃ/ is transcribed as /ɕ/ by some linguists (Pan et al, 2004). As for /s/, as indicated by Beckman, Edward, and Li (2009), the English /s/ is clearly alveolar while the Japanese /s/ is more laminal and possibly somewhat dentalized. Indeed, knowledge of these cross-linguistic phonetic differences would contribute to attainment of native-like accents, particularly for

[3] The original phonetic value of the consonant in the modern /s/ column in the *kana* chart is still under debate, but there is evidence that it was a coronal affricate rather than a fricative (Takayama, 2003).

[4] For *kana* typing, pressing "s-w-i" or "s-u-x-i" makes <すぃ> that corresponds to [si].

advanced level learners who can afford to deal with articulatory details. However, this present study does not discuss phonetic details.

## 2 Experiment 1: Reading task

### 2.1 Methodology and stimuli

42 Japanese ESL learners were recorded reading aloud unfamiliar words <sith>[5] and <shiff>, which are expected to be pronounced as [sɪθ] and [ʃɪf] based on the English GPC. I used nonsense words in order to observe the participants' pure GPC knowledge without loanword interference. I chose <i> as the following vowel letter rather than <ee> (corresponding to /iʲ/) since the potentially confusing Japanese romanization spellings are <si> and <shi> as shown in Figure 1. In addition, <ee> could be pronounced as [ɛ] or similar variations, rather than [iʲ], due to Japanese romanization interference. In such cases, [s, ʃ] before high front vowels cannot be observed. As for the coda consonants, I chose relatively difficult sounds for Japanese speakers (Japanese /f/ is bilabial [ɸ] rather than [f], and Japanese lacks /θ/) in order to draw their attention to the coda consonants. These stimuli were mixed with those for another study in which I examine Japanese ESL learners' knowledge of English vowel spellings, so the participants pronounced 50 nonsense words in total. In this present study, I analyzed only these two words. All the stimuli were printed on a sheet of paper.

### 2.2 Participants

In total, 42 Japanese ESL learners were recruited in Victoria, British Columbia in Canada. They were divided into two groups: 1) ESL learners who have been in Canada for 1 year or more, and 2) Japanese ESL learners who have been in Canada for less than 1 year. In the first (longer length of residence in Canada (LOR)) group, there were 26 participants (10 males and 16 females). Their mean LOR was 5.2 years (ranging from 1 year to 20 years). Their mean age was 33.9 years old (ranging from 19 to 71). In the second (shorter LOR) group, there were 16 participants (3 males and 13 females). Their mean LOR was 5.7 months (ranging from 3 weeks to 11 months). Their mean age was 25.1 years old (ranging from 19 to 32). As a control group, I also recruited 14 native speakers of Canadian English (8 males and 6 females) at the mean age of 31.9 (ranging from 20 to 56). No participants reported a hearing problem.

As limitations, I originally planned to compare English-as-a-foreign-language (EFL) learners, or inexperienced learners, and English-as-a-second-language learners, or experienced learners, but due to my physical presence in Canada and the time constraint, I collected data only from Japanese ESL learners. Thus, length of residence and their English proficiency could not be controlled.

[5] I intended to make nonsense words, but <sith> turned out to be an existing word, *Sith* (an organization in Star Wars).

### 2.3 Results of the reading task

Table 1 shows the results from Japanese ESL participants whose LOR is 1 year or more. Table 2 shows the results from those whose LOR is less than 1 year.

| Written stimulus | Productions [s] | [ʃ] | [st] | Total |
|---|---|---|---|---|
| <sith> | 25 | 1 | | 26 |
| <shiff> | 8 | 18 | 1 | 26 |

*Table 1: Results of the reading task from the longer LOR group*

| Written stimulus | Productions [s] | [ʃ] | [st] | Total |
|---|---|---|---|---|
| <sith> | 16 | | | 16 |
| <shiff> | 12 | 3 | 1 | 16 |

*Table 2: Results of the reading task from the shorter LOR group*

Surprisingly, only 1 participant, whose LOR is 5 years, out of the total 42 participants, pronounced <s> as [ʃ]. She was first going to pronounce [s] and then changed it to [ʃ], suggesting that she had a dilemma but chose to pronounce [ʃ] over [s]. All the other participants correctly pronounced <s> before <i> as [s], although 4 in the longer LOR group and 3 in the shorter LOR group pronounced <i> as [aj], meaning that not all the participants pronounced [s] before a high vowel. All the other participants' vowels were in the high front region like [ɪ] or [i]. One participant of each group pronounced <sh> as [st], presumably because they mistook <shiff> for the real word <stiff>. Other than these 3 errors (<s>→[ʃ] × 1, <sh>→[st] × 2), all the other 20 errors were <sh> pronounced as [s]. These results turned out completely opposite to the popular assumption that Japanese ESL learners are unable to produce [s] before high front vowels and tend to substitute /s/ with [ʃ]. Rather, Japanese ESL learners are strongly biased towards [s]. As shown in Figure 1, the <shi>-[si] correspondence does not exist in Japanese romanization, so their <sh>→[s] errors cannot be L1 interference. Rather, they may have assumed that [ʃ] sounds more Japanese and hyper-corrected it to [s]. This pattern is consistent with Eckman and Iverson (2013), who found that Japanese ESL learners produced a preponderance of hypercorrection errors, that is, /ʃ/ before high front vowels as in *sheep* pronounced as [s]. As for LOR and error frequency, it is noteworthy that in the shorter LOR group, 12 out of 16 (75%) of the participants pronounced <sh> as [s], while in the longer LOR group, only 8 out of 26 (31%) pronounced <sh> as [s]. This implies that this hypercorrection becomes less frequent as they are exposed in English for longer time.

Interestingly, the results of this present study (as well as Eckman and Iverson's (2013) study) contradict those in my previous study (Nogita, 2010). In Nogita's (2010) passage reading task by beginner to lower-intermediate Japanese ESL learners, <s> in real English words (*see, sea, sits, seat, CD*, and *sick*) was

pronounced as [ʃ] 21% of the time, and <sh> in real words (*washing*, *relationship*, *sheets*, and *she*) was pronounced as [s] 19% of the time, meaning that <s> and <sh> were almost equally frequently mispronounced. However, this discrepancy can be explained by error frequency differences among words in Nogita's (2010) experiment. Based on One-Way ANOVA, <s>(and <c>)-related error frequency significantly differs depending on words (from 0% for *see* to 33% for *CD/sits*, $F(5, 156) = 3.61$, $p < 0.005$), and <sh>-related error frequency marginally significantly differs (from 11% for *washing* to 37% for *sheets*, $F(3, 104) = 2.34$, $p = 0.078$). High error frequency in at least *CD* (33%) and *seat* (30%) can be analyzed as loanword interference (c.f. [ʃi:di:] 'CD' and [ʃi:to] 'seat' in Japanese), suggesting that Japanese ESL learners may be more likely to mispronounce <s> as [ʃ] in English words that have been already imported to the Japanese vocabulary. In contrast, in unfamiliar words, Japanese ESL learners are very unlikely to mispronounce <s> as [ʃ], even with short LOR.

Japanese ESL learners' <sh>→[s] error pattern may be analogous to English speaking people's <j>→[ʒ] pattern in the Chinese loanword *Bei*j*ing*, which is called "hyperforeignization" (Janda, Joseph, & Jacobs, 1994, 71). The original Chinese sound of <j> in *Bei*j*ing* is the alveo-palatal affricate [dʑ], which is more similar to the English [dʒ] than to [ʒ] (Janda at al., 1994, 80). In addition, based on the basic English GPC rule, the letter <j> commonly corresponds to /dʒ/. Moreover, English has the phonological /dʒ/-/ʒ/ contrast as in *J*ap*an* and *A*s*ia*, so this distinction should not be a problem for English speaking people. Despite all of these legitimate reasons for choosing [dʒ], English speaking people have selected [ʒ], since English speakers tend to treat palato-alveolar fricatives /ʒ, ʃ/ as generic foreign consonants (Janda at al., 1994). Likewise, in Japanese, [s] before high front vowels sounds foreign, so Japanese ESL learners may choose [s] for <sh> by prioritising foreignness over the L1 spelling rule and the original L2 sound.

Finally, as for the control group, unsurprisingly, all the 14 native English speakers pronounced <sith> as [sɪθ] and <shiff> as [ʃɪf] without hesitation.

## 3 Experiment 2: Spelling task

### 3.1 Methodology and stimuli

In the spelling task, the same participants were asked to listen to the audio stimuli [siʲv] and [ʃiʲʧ] and spell what they heard. They were allowed to listen to the stimuli as many times as they wanted. These audio stimuli were pronounced by a phonetically trained male native Canadian English speaker from British Columbia in his 40's. I selected the tense /iʲ/ as in *eat* as the following vowel, but not the lax /ɪ/ as in *it,* since the English lax /ɪ/ can be perceived as the Japanese /e/ by Japanese L1 speakers. In Japanese romanization, /se/ and /ʃe/ are clearly spelled differently, <se> and <she> (or <sye>) respectively, so the participants' <s>-<sh> confusion would not be expected if they perceive the English /ɪ/ as the Japanese /e/. Again, these two stimuli were mixed with those in the study of

vowel spelling, so the participants listened to many more stimuli other than [siʲv] and [ʃiʲɡ].

### 3.2 Results of the spelling task

Table 3 shows the results from Japanese ESL participants whose LOR is 1 year or more. Table 4 shows the results from those whose LOR is less than 1 year.

| Sound stimulus | Spellings <s> | <sh> | <c> | <th> | <ch> | Total |
|---|---|---|---|---|---|---|
| [siʲv] | 14 | 3 | 3 | 6 | | 26 |
| [ʃiʲɡ] | 8 | 16 | | 1 | 1 | 26 |

*Table 3: Results of the spelling task from the longer LOR group*

| Sound stimulus | Spellings <s> | <sh> | <c> | <th> | <sch> | Total |
|---|---|---|---|---|---|---|
| [siʲv] | 10 | 3 | 2 | 1 | | 16 |
| [ʃiʲɡ] | 6 | 8 | | 1 | 1 | 16 |

*Table 4: Results of the spelling task from the shorter LOR group*

In the spelling task, Japanese ESL participants in both LOR groups more frequently spelled [ʃ] as <s> (14/42, 33%) than [s] as <sh> (6/42, 14%), suggesting that they were more biased towards <s> than towards <sh>. As shown in Figure 1, in Japanese romanization, the sound [ʃi] can be spelled as <si>, so their English [ʃ]→<s> pattern is likely L1 interference. It should be noted that as shown in Appendix A, many participants spelled the following vowel [iʲ] as <ee>, <ea> and other variations involving <e>, rather than Japanese romanization-like <ii>, <ih> or others starting with <i>. This suggests that many of the participants had acquired the English vowel phoneme-to-grapheme conversion /iʲ/→<ee, ea> to some extent, but had not acquired the consonant phoneme-to-grapheme conversion /ʃ/→<sh> before /iʲ/. In other words, their [ʃiʲ]→<see, sea> conversion patterns are partial L1 transfer. As for the opposite type of errors, specifically [s] spelled as <sh>, the [si]-<shi> correspondence does not exist in Japanese romanization as shown in Figure 1, so this pattern should not be L1 interference. This may be another type of hypercorrection. As for the frequency of the L1-influence-type errors ([ʃ] spelled as <s>), it was 6 out of 16 (38%) in the shorter LOR group, whereas it was 8 out of 26 (31%) in the longer LOR group, meaning that L1 influence slightly reduces but not dramatically. The frequency of the hypercorrection-type errors ([s] spelled as <sh>) was 3 out of 16 (19%) in the shorter LOR group, whereas it was 3 out of 26 (12%) in the longer LOR group, meaning that hypercorrection also slightly reduces but not dramatically. About individual patterns, only one participant in the shorter LOR group (11 months) exactly oppositely spelled [s] as <sh> and [ʃ] as <s>, but all the other participants' error patterns were biased towards either <s> or <sh>.

Notice that there were more variations of the participants' responses in the spelling task than in the reading task. The sound [s] spelled as <c> is regarded as correct as long as <c> precedes <e> or <i>. The [s]-<c> correspondence does not exist in Japanese romanization, indicating that those participants seem to have acquired the English so-called Soft-C ([s]-<c>) rule at least to some extent. Other patterns are [ʃ] spelled as <sch> or <ch>. These correspondences do not exist in Japanese romanization, so these should not be L1 transfer. Indeed, these patterns exist in English words, as in *chef* [ʃɛf] and *schedule* [ʃɛʤu$^{w}$l] in some dialects, so <sch, ch> for [ʃ] can be regarded as acquisition of less common L2 GPC patterns rather than acquisition of the default GPC patterns.

Finally, L1-L2 phonology-related errors are [s] or [ʃ] spelled as <th>. Although they did not hear [θ], they were biased towards <th>, which is another type of hypercorrection. Under the assumption that they know the English /θ/-<th> GPC rule, the fact that even by participants in the longer LOR group, [s] was spelled as <th> by 6 participants and [ʃ] was spelled as <th> by 1 participant implies difficulty in acquisition of non-L1 phoneme /θ/ by Japanese ESL learners.

As for the control group, unsurprisingly, all the 14 native English speakers spelled /s/ as <s> and /ʃ/ as <sh> although spellings of the rhyme parts showed a few variations.

## 4 Comparison between reading and spelling tasks

What is interesting is that not all the participants who made an error in the reading task also made an error in the spelling task, and vice versa. For example, those who read the grapheme <sh> as [s] did not necessarily spell the sound [s] as <sh>, and those who spelled [s] as <sh> did not necessarily read <sh> as [s]. Table 5 shows their error patterns. "Wrong in reading" indicates the number of participants who made an <s>-<sh>-related error in the reading task; note that other errors (e.g. <sh>→[st]) are not included. "Wrong in spelling" indicates the number of participants who made a [s]-[ʃ]-related error in the reading task; other errors (e.g. [s]→<th>) are not included. "Wrong in one task" indicates that the number of participants who made a [s]/<s>-[ʃ]/<sh>-related error in one task but did not in the other task. "Wrong in both tasks" indicates the number of participants who made a [s]/<s>-[ʃ]/<sh>-related error in both tasks. "Correct in both tasks" indicates the number of participants who did not make any [s]/<s>-[ʃ]/<sh>-related errors, and bracketed numbers show the number of those who did not make any other type of errors as well, such as [s, ʃ]→<th>. "1+ year" indicates the group with a 1 year or more of LOR and "< 1 year" indicates the group with a less than 1 year of LOR. The numbers "9/26", for example, indicates 9 out of 26.

| | Wrong in reading | Wrong in spelling | Wrong in one task | Wrong in both tasks | Correct in both tasks |
|---|---|---|---|---|---|
| 1+ years | 9/26 | 11/26 | 12/26 | 4/26 | 10(7)/26 |
| < 1 year | 12/16 | 8/16 | 6/16 | 7/16 | 3(1)/16 |
| Total | 21/42 | 19/42 | 18/42 | 11/42 | 13(8)/42 |

*Table 5: Error patterns in both tasks*

Interestingly, in total, 18 out of 42 (43%) participants made at least one error in one of the reading or spelling tasks but not in the other task, meaning that their reading patterns and spelling patterns were inconsistent. Among the 11 (26%) participants who made errors in both tasks, 1 in the longer LOR group and 2 in the shorter LOR group pronounced [s] for both <s> and <sh> in the reading task, but spelled <sh> for both [s] and [ʃ] in the spelling task. The other 8 pronounced [s] for both <s> and <sh> in the reading task and spelled <s> for both [s] and [ʃ] in the spelling task.

All this indicates that these 29 out of 42 (69%) participants' grapheme-to-phoneme conversion and phoneme-to-grapheme conversion patterns were inconsistent. There is a possibility that at least some of these 29 participants' error patterns were random or at the pre-systematic stage. Only 8 out of 42 (19%) participants did not make any errors, including errors like [s] spelled as <th>. Interestingly, while the one whose LOR was only 5 months did both tasks correctly, the one whose LOR was the second longest (14 years) made an error in both tasks, suggesting that although there is a tendency that those in the longer LOR group performed better, LOR does not guarantee their acquisition of <s>-/s/ and <sh>-/ʃ/ correspondence. Finally, reportedly, 4 were or had been majoring in linguistics, but only 1 of them did both tasks correctly, implying that phonetics and IPA knowledge does not guarantee their acquisition of <s>-/s/ and <sh>-/ʃ/ correspondence, although there needs to be more careful research about the relationship between English GPC knowledge and phonetic knowledge to make a conclusion.

## 5 Discussion

### 5.1 Teaching implication

According to my impression, quite a few Japanese ESL learners feel that they have difficulty in the English /s/-/ʃ/ contrast. However, Nogita (2010) already proved that Japanese does have the /si/ and /ʃi/ contrast and that even monolingual Japanese speakers can easily distinguish /si/ and /ʃi/ if these are in Japanese contexts. To help those Japanese ESL learners, first, it is very important to let them aware that both /si(:)/ and /ʃi(:)/ do exist in Japanese contrastively, as in しいたけ /ʃi:take/ 'shiitake mushroom' and まつばらスィー /matsubarasi:/ (a

handle of a Japanese person) (retrieved on Jan. 31, 2016 from https://www.youtube.com/watch?v=jE2UzFRAADc). Then, have them aware that the English /s/-/ʃ/ confusion is only a spelling issue, that is, most of the time <s> corresponds to /s/ and <sh> corresponds to /ʃ/. It would also be helpful to let them know the potentially confusing L1 spelling rule (both *cabinet-ordered rōmaji* <si> and *Hepburn rōmaji* <shi> corresponding to /ʃi/) and the common hypercorrection (the English <shi> pronounced as [sɪ] as in *friendship*). After they fully acquire these spelling rules and reach the level at which they can afford to deal with articulatory details without being confused, articulation training regarding cross-linguistic phonetic differences in /s/ and /ʃ/ can be given if necessary.

### 5.2 Limitation

Since this study was done only with ESL learners, it is unknown whether my conclusion can be generalized for Japanese EFL learners. In addition, the number of stimuli was few. This is because this experiment was done together with a vowel study as mentioned above, and I avoided making too many stimuli in order to reduce participants' loads. There is also a possibility that at least a few of the spelling errors may have been due to mishearing. For example, [ʃ] spelled as <ch> may have been mishearing of [ʃ] as /ʧ/ (although Japanese does have the /ʃ/-/ʧ/ contrast). This experiment could not distinguish pure spelling errors and potential mishearing.

## 6 Conclusion

The present study supports the argument that Japanese ESL learners confuse /s/ and /ʃ/ before high front vowels in English because of their lack of knowledge of the English grapheme-phoneme correspondence rules, not because of difficulty in articulation. It is likely that their confusion originates from the complex Japanese romanization spelling patterns as shown in Figure 1. In the reading task, contrary to the popular assumption that Japanese ESL learners tend to substitute [ʃ] for /s/ due to the L1 phonotactic constraint, the results showed their strong tendency of pronouncing <sh> as [s], and only 1 out of 42 participants pronounced <s> as [ʃ]. In the spelling task, although the tendency was not as clear as the reading task, they tended to spell [ʃ] as <s> more frequently than [s] as <sh>, presumably due to L1 Japanese romanization interference. Most importantly, 29 out of 42 (69%) participants' grapheme-to-phoneme conversion and phoneme-to-grapheme conversion patterns were not consistent, suggesting that they had not acquired the English default GPC rules, <s>-/s/ and <sh>-[ʃ].

**Acknowledgements**

I am grateful to Dr. Hua Lin and Dr. Sonya Bird for their supervision. I am also very grateful to all the 56 participants of the experiments. Also, I am thankful for the reviewers and the WPLC staff.

**References**

Avery, P. & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford University Press.

Beckman, M. E., Edwards, J., & Li, F. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journals of Phonetics, 37*, 111-124.

Department of English Language, College of Arts and Sciences, The University of Tōkyō, Komaba (2009). *Nihongo no Rōmajihyōki no Suoshōkeishiki [recommended rōmaji spelling]*

Eckman, F., & Iverson, G. K. (2013). The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition*, *35*(01), 67-92.

Janda, R. D., Joseph, B. D., & Jacobs, N. G. (1994). Systematic hyperforeignisms as maximally external evidence for linguistic rules. *The reality of linguistic rules*, *26*, 67-92.

Matsuzaki, H. (1993). A quantitative study of fluctuation in Japanese loanword-phone notation. *Journal of the Department of Japanese, Tohoku University, 3*, 3-94.

Ministry of Education, Culture, Sports, Science and Technology-Japan (2009), *Rōmaji no Tsuzurikata [rōmaji spelling]*, Retrieved January 28, 2016, from http://www.mext.go.jp/b_menu/hakusho/nc/k19541209001/k19541209001.html

Nogita, A. (2010). *Examination of the [si] and [ʃi] Confusion by Japanese ESL Learners* (Master's Thesis, University of Victoria).

Pan, H. Y., Utsugi, A., & Yamazaki, S. (2004). An acoustic phonetic study of voiceless alveolo-palatal fricatives in Japanese, Korean and Chinese. *Journal of General Linguistics, 7*, 1-27.

Takayama, M. (2003). Issues in Japanese historical phonology. *Journal of the Phonetic Society of Japan*, *7*(1), 35-46.

Taylor, I., & Taylor, M. M. (1995). *Writing and literacy in Chinese, Korean and Japanese* (Vol. 3). John Benjamins Publishing.

Vance, T. J. (2008). *The sound of Japanese.* New York: Cambridge University Press.

**Appendix A (Japanese ESL participants' raw data)**

The table below shows each Japanese ESL participant's responses in both reading task and spelling task in both groups along with reported Length Of Residence in Canada (LOR). Capitalizations are based on their original spellings.

| Shorter LOR ESL group | | | | | Longer LOR ESL group | | | | |
|---|---|---|---|---|---|---|---|---|---|
| LOR | Stimuli | | | | LOR | Stimuli | | | |
| | Reading task | | Spelling task | | | Reading task | | Spelling task | |
| month | <sith> | <shiff> | [siʲv] | [ʃiʲʧ] | year | <sith> | <shiff> | [siʲv] | [ʃiʲʧ] |
| 0.7 | [s] | [ʃ] | seeb | thezu | 1 | [s] | [ʃ] | seeph | sheech |
| 2 | [s] | [s] | ceeb | seeche | 1 | [s] | [ʃ] | thiegh | sitch |
| 2 | [s] | [s] | seeve | shiech | 1.3 | [s(aj)] | [s] | seave | sechu |
| 3 | [s] | [st] | seve | schitch | 1.5 | [s] | [ʃ] | Thieve | Seech |
| 3 | [s(aj)] | [s] | sheeb | sheach | 1.5 | [s] | [s] | ceive | seazue |
| 3 | [s] | [ʃ] | sheeb | shirtch | 1.7 | [s] | [ʃ] | seave | chuich |
| 4 | [s] | [s] | seave | seach | 2 | [s] | [ʃ] | seeb | seech |
| 5 | [s] | [ʃ] | seaf | sheech | 2 | [s(aj)] | [ʃ] | theeve | shesh |
| 5 | [s] | [s] | ceive | seech | 2.9 | [s] | [st] | theaf | siechi |
| 5 | [s] | [s] | seave | sheech | 2.9 | [s] | [ʃ] | seaf | seatue |
| 8 | [s] | [s] | theef | sheech | 3 | [s] | [s] | seeve | shech |
| 9 | [s] | [s] | seef | seatch | 3 | [s(aj)] | [ʃ] | seeve | sheech |
| 10 | [s] | [s] | seab | sheech | 4.5 | [s] | [s] | seeve | shityu |
| 10 | [s] | [s] | seeb | seech | 5 | [s] | [ʃ] | shif | shich |
| 11 | [s] | [s] | seeve | shitu | 5 | [ʃ] | [ʃ] | Seev | Shisu |
| 11 | [s] | [s] | shive | sitw | 5.3 | [s] | [s] | seef | seech |
| | | | | | 6 | [s] | [s] | thieve | thiech |
| | | | | | 6 | [s] | [ʃ] | seef | sheech |
| | | | | | 7 | [s] | [ʃ] | sheeve | shetch |
| | | | | | 7 | [s] | [ʃ] | seeve | sheech |
| | | | | | 8 | [s] | [ʃ] | seave | shich |
| | | | | | 8 | [s(aj)] | [ʃ] | theib | shich |
| | | | | | 10 | [s] | [ʃ] | cib | sheech |
| | | | | | 10 | [s] | [s] | seeb | sheich |
| | | | | | 14 | [s] | [s] | sheev | sheech |
| | | | | | 20 | [s] | [ʃ] | ceive | sheich |

**Appendix B (Canadian English participants' raw data)**

The table below shows native Canadian English speakers responses in the spelling task. Capitalizations are based on their original spellings.

| Stimuli | |
|---|---|
| [siʲv] | [ʃiʲʧ] |
| seeve | sheech |
| seeve | sheech |
| Seev | shech |
| Sieve | Sheech |
| seeve | sheech |
| seeve | sheech |
| Seeve | Sheech |
| seeve | sheech |
| SEEV | SHEECH |
| seeve | sheetch |
| seeve | sheech |
| seeve | sheech |
| seethe | sheech |
| seeve | sheech |

# Linguistic variation 'from below': Northern German in nineteenth-century North Frisia

**Timothy C. Jacob-Owens**
University of Bristol
*tj1066@bristol.ac.uk*

This paper presents a linguistic analysis of a newly-compiled corpus of historical correspondence. The corpus comprises 18 private letters written in High German by North Frisians between 1839 and 1851. The investigation is thus conceived in the spirit of 'language history from below', a highly topical approach within the field of historical sociolinguistics. The paper seeks to identify commonalities between the variety of High German written by North Frisians in private correspondence and the varieties of High German spoken elsewhere in northern Germany. The letters are analysed for evidence of non-standard diatopically-marked linguistic variation in the realms of phonology and morpho-syntax. This analysis demonstrates that the letters exhibit a number of classic northern German phonological features, such as *g*-spirantisation and affricate reduction. Several northern German morpho-syntactic structures are also attested. The results thus suggest that the informal writing of North Frisians shared a number of common linguistic traits with regional varieties of High German used elsewhere in northern Germany.
*Keywords: historical sociolinguistics; diatopic variation; nineteenth century; North Frisia*

## 1 Introduction

Traditional language histories are conceived (mono-)linguistically rather than geographically. Hence, a history of the German language is not a linguistic history of Germany. Moreover, such studies tend to focus solely on the standard language of a given era, such as Standard High German, i.e. the language of the educated élite (cf. Milroy, 2012). For these reasons, the role of regional (diatopic) and/or social (diastratic) factors in determining linguistic variation has, until recently, remained conspicuously absent from historical linguistics. Similarly, the traditional focus on formal, printed text sources has hindered research on the nature of the spoken language of the past. Over the past decade, historical sociolinguists have, however, made great progress in addressing these desiderata. For instance, Elspaß (2005) demonstrated through an analysis of nineteenth-century German emigrants' letters that the range of diatopically- and diastratically-marked variation was far greater than suggested by the

contemporary standard print norm. In a similar vein, though considerably more modest in scope, the present paper explores non-standard diatopically-marked linguistic variation in a newly-compiled corpus of nineteenth-century letters written in High German by North Frisians. The recently published first volume of the *Norddeutscher Sprachatlas* (Elmentaler & Rosenberg, 2015) has shown that a wide range of diatopically-marked phonological variants are shared across a large area of northern Germany today. The extent to which this was the case historically is, however, still poorly understood. While diatopic variation elsewhere in nineteenth-century northern Germany has already been explored (cf. Elspaß, 2005; Langer, 2013), the High German written (and spoken) in nineteenth-century North Frisia has not yet been investigated. This paper aims to address this desideratum and, in particular, seeks to identify commonalities between the variety of High German written (and spoken) by North Frisians and the varieties of High German spoken elsewhere in northern Germany. The paper is structured as follows: the first section sketches the historical-sociolinguistic context of the investigation; the second section discusses the study's research method; the third section presents a linguistic analysis of the data in line with the paper's aims; and the final section provides some brief concluding remarks.

## 2 The historical-sociolinguistic context

North Frisia is the westernmost region within the German-Danish borderlands, comprising its western coastal region and the islands of Föhr, Amrum, Sylt, and the Halligen (see map, below). In common with the rest of the German-Danish border region, North Frisia has for centuries been characterised by multilingualism. In addition to the region's non-dominant autochthonous language (cf. Clyne, 1992), North Frisian, Low German was spoken in North Frisia as the dominant lingua franca and language of prestige from the fourteenth to the early eighteenth century (Faltings, 1992, pp. 54-55). With the rise of nationalism in the nineteenth century, standardised 'national' languages were frequently politicised as markers of national identity. In this vein, the nineteenth- century German-Danish national conflicts saw the introduction of aggressive language policy measures targeting schools, the church, and public administration, which aimed either to Germanise or Danicise the inhabitants (cf. Langer, 2014). In the context of North Frisia, this meant the imposition of High German as the dominant language of schooling and religion (Jensen, 1961, p.
256). Anecdotal evidence indicates that North Frisian, and in some cases Low German, remained North Frisians' mother tongue(s), and that High German was limited to those domains in which it was imposed 'from above' (Kohl, 1846; Jensen, 1961). This raises the interesting question of how the High German written and spoken in North Frisia may have been influenced by the other languages present at the time, i.e. North Frisian, Low German and other regional varieties of High German. We may wonder to what extent the High German spoken by North Frisians shared linguistic features with the High German spoken elsewhere in northern Germany, if indeed their principal and perhaps only

contact with the language was through standardised High German sources such as schoolbooks and the Bible.

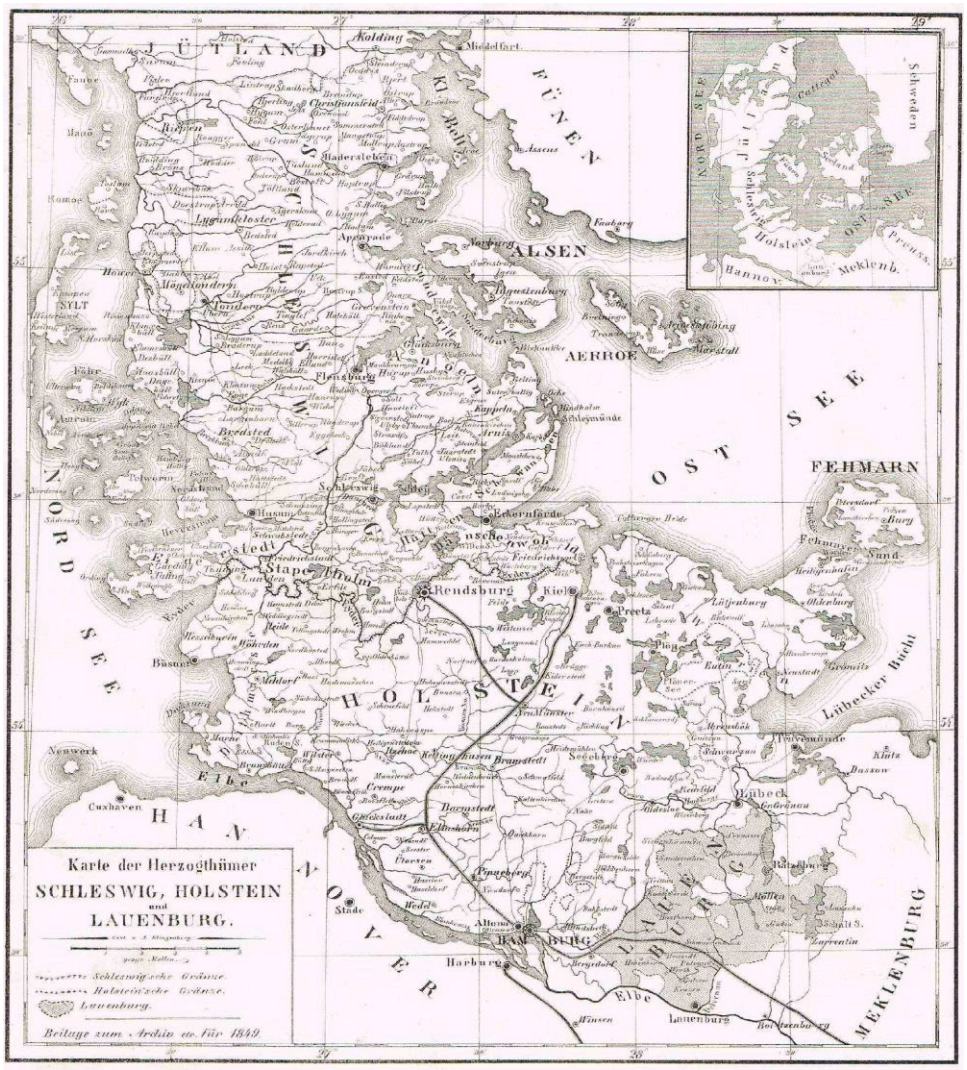



Figure 1. *Map of the German-Danish Border Region 1849*. Retrieved from https://www.spsh.uni-kiel.de/

## 3 Methodology

### 3.1 Linguistic variation 'from below'

The present paper is conceived within the methodological framework of 'language history from below', a highly topical approach within historical sociolinguistics. As Labov (1994, p. 11) famously observed, historical linguists must make do with 'bad data', i.e. incomplete records that survive by random

chance and that often do not accurately represent the spoken language of their authors. For historical *socio*linguists this problem is even more acute, given the scarcity of historical written data which exhibit diatopically- and/or diastratically-marked variation. In recent years, historical sociolinguists have approached language history 'from below', sourcing data from so-called ego-documents (e.g. private letters and personal journals) written by less-educated and/or lower-class individuals as a means by which to glean insights into the linguistic situation 'below' (cf. Elspaß, 2005). This approach is not only useful because it provides insights into the language of the low- and middle classes, but also because it presents an unparalleled picture of the spoken language of the past. In discussing the range of possible sources of linguistic data, Koch & Oesterreicher (1990, p. 5) identify and differentiate two types of language *medium*, phonic and graphic, and two types of *conception*, namely written and oral. Within this framework, ego-documents are presumed to be graphic realisations of conceptual orality. Thus, the language of private letters and journals is widely taken to be 'as close to speech as non-fictional historical texts can possibly be' (Rutten & van der Wal, 2014, p. 4).

### 3.2 Corpus description

The present paper draws on a newly-compiled corpus (henceforth ÖA corpus) of nineteenth-century correspondence found in the North Frisian Öömrang Archiif (ÖA) at the Ferring Stiftung (Alkersum auf Föhr, Germany). The ÖA corpus comprises 18 private letters (approximately 10,000 tokens) written between 1839 and 1851 to Knudt Jungbohn Clement (henceforth KJC), a North Frisian historian, by members of his family, all of whom came from the North Frisian island of Amrum. Four of the letters were written by KJC's brother Olde Jung Clement (OJC), four by his brother Boy Olde Clement (BOC), four by his brother-in-law Wellam Peters (WP), three by his mother Kerrin Hansen (KH), two by his sister Keike Wellams (KW), and one by his stepfather Hans Sponagel (HS). The following analysis draws on data from all six writers.

A key methodological consideration within the 'language history from below' approach is the so-called 'writer-sender problem', i.e. the question of whether the writer and the sender of a given letter are one and the same person (Nobels & van der Wal, 2012, p. 348). This is, of course, crucial to linking linguistic traits to particular biographical details, such as place of birth or level of education. In the context of the present study, the writer-sender problem has been resolved by following the Leiden Identification Procedure (LIP), which identifies a number of characteristics which indicate whether a given letter is a genuine autograph. The first of these are so-called 'content clues', i.e. metalinguistic comments within a letter that point to the named writer being the genuine author (Nobels & van der Wal, 2012, p. 349). The ÖA corpus contains a variety of content clues in a number of different letters. For instance, KH, who wrote her letters between the ages of 62 and 73 and who thus might reasonably be expected to have had someone write them on her behalf, writes in one letter 'I wanted to

write more, but my thoughts are at an end' ('*ich wolte noch wohl was mehr schreiben aber meinen Gedanken sind zu kortz*'), strongly suggesting that she herself is the author of the letter. Another LIP indicator of a letter's autograph status is a handwriting match between two letters signed by the same person (Nobels & van der Wal, 2012, p. 350). In the case of the ÖA corpus, each author has an individual handwriting style and signature, both of which remain consistent in all of their letters (for examples, see the project website https://www.spsh.uni-kiel.de/archiv). Hence, we may be reasonably certain that the letters are genuine autographs. The data thus represent an authentic sample of the High German written (and spoken) on Amrum between 1839 and 1851.

### 3.3 Analytical approach

The analysis in this paper focuses on northern German phonological variants and a series of non-standard grammatical constructions. The latter of these include non-standard case morphology, dative nominal possessive constructions, split pronominal adverbs, and non-standard use of *wie* and *wo*. These phenomena have all been attested in previous studies of nineteenth-century northern German private writings (e.g. Elspaß, 2005; Denkler & Elspaß, 2007; Langer, 2013). Close qualitative analysis is required in order to precisely characterise these features and, where relevant, to identify possible correspondences in Low German and/or North Frisian.

## 4 Analysis

### 4.1 Phonology

#### *4.1.1 g-Spirantisation*

*g*-Spirantisation, i.e. the realisation of /g/ as a fricative, is a characteristic feature of northern High German: *Tag* ('day') is pronounced /taç/ in the North and /ta:k/ in the South. Today, while *g*-spirantisation in word- or morpheme-initial position is largely restricted to north-eastern Germany (e.g. Berlin and Brandenburg), it is commonly found in word- or morpheme-final position across the whole of northern Germany (Elmentaler & Rosenberg, 2015, pp. 237-239, 261-263). In written High German, the fricative /ç/ is typically represented orthographically as <ch> and one might thus expect to find examples in the ÖA corpus where <ch> appears in place of Standard German (SG) <g>, e.g. '*wechgekommen*' for SG *weggekommen* (Langer, 2013, p. 82). While no such examples appear in the ÖA corpus, the letters do contain a number of examples of hypercorrection, i.e. <g> in place of SG <ch>, demonstrating the writers' awareness of this particular northern feature. As expected, all the examples appear in either morpheme- (examples 1 and 2) or word-final (example 3) position. As Denkler & Elspaß (2007, p. 93) discuss, hypercorrection of this sort is 'typical of non-dialectal varieties with regional colouring' ('*typisch für nicht-dialektale Varietäten mit*

*regionaler Färbung*'). Similar examples are also attested in Langer's (2013, p. 82) corpus of nineteenth-century northern German soldiers' letters.

(1) BOC: *beträgtliches*
SG: *beträchtliches* /bətrɛçtlıçəs/
'considerable'

(2) HS: *Quiting büger*
SG: *Quittungsbücher* /kvıtʊŋsby:çɐ/
'receipt books'

(3) HS: *reichlig*
SG: *reichlich* /raıçlıç/
'amply'

*4.1.2 t-Apocope*

Another common trait of northern High German is *t*-apocope, i.e. loss of word-final /t/, e.g. the realisation of *nicht* ('not') as /nıç/ rather than SG /nıçt/. Today, this feature is prevalent across northern Germany (Elmentaler & Rosenberg, 2015, p. 277). Examples of this diatopically-marked feature are attested in two different nineteenth-century letter corpora by Langer (2013, p. 83) and Elspaß (2005, p. 440), the latter of whom identifies the phenomenon as a 'characteristic of northern German texts' ('*Merkmal norddeutscher Texte*', Elspaß, 2005, p. 456) from the nineteenth century. In the ÖA corpus, *t*-apocope does not occur in commonly-affected monosyllabic lexemes such as *nicht*, *jetzt* and *ist* (cf. Elspaß, 2005, p. 440). This points to a high level of awareness of the standard norms on the part of the writers. In fact, the ÖA corpus contains just one token of *t*-apocope, namely in the letter by HS (see example 4). No examples of the corresponding SG form occur in this letter.

(4) HS: *abgesetz* /apgəzɛts/
SG: *abgesetzt* /apgəzɛtst/
'discontinued'

*4.1.3 Affricate reduction*

A further characteristic non-standard feature of northern High German is the realisation of the affricates /ts/ and /pf/ as the fricatives /s/ and /f/ respectively (Elmentaler & Rosenberg, 2015, p. 291). This can be explained as a kind of interference feature from Low German, which does not have the affricates /ts/ and /pf/ having not undergone the Second Sound Shift (Elmentaler & Rosenberg, 2015, p. 291). Today, the feature is found across the whole of northern Germany (Elmentaler & Rosenberg, 2015, p. 293, 297). The ÖA corpus contains a single example of the reduction of /ts/ to /s/, namely '*su*' (see example 5). The

corresponding SG form *zu* occurs a total of 82 times. Similar examples are attested by Langer (2013, p. 83). The ÖA corpus also contains instances of hypercorrect <z>, the grapheme used to represent /ts/ orthographically, in place of <s> (see examples 6 and 7). One of these examples, namely '*zchlechte*' (see example 6), was subsequently corrected by the author to SG *schlechte*, suggesting a strong metalinguistic awareness of this non-standard variant. The ÖA corpus features just one example of the reduction of the affricate /pf/ to the fricative /f/, which appears in a letter by KH (see example 8). No examples of the corresponding SG form *Strümpfe* occur in the letters of KH or in the rest of the ÖA corpus. This variant is not attested in Elspaß (2005) or Langer (2013).

(5) KH: *su* /suː/
SG: *zu* /tsuː/
'to'

(6) KH: *zchlechte*
SG: *schlechte* /ʃlɛçtə/
'bad'

(7) WP: *Kozt*
SG: *Kost* /koːst/
'cost'

(8) KH: *Strümfe* /ʃtʀymfə/
SG: *Strümpfe* /ʃtʀympfə/
'stockings'

*4.1.4 Unvoiced /s/*

A less common non-standard phonological feature of northern High German is the realisation of <s> as unvoiced /s/ rather than SG voiced /z/ (Elmentaler & Rosenberg, 2015, p. 334). Today, unvoiced /s/ in word-initial position is particularly prevalent in North Frisia (Elmentaler & Rosenberg, 2015, p. 331). Unvoiced /s/, represented orthographically as either <ss> or <ß>, occurs most frequently in the letters of KH, which contain four tokens of '*Hausse*' (see example 9), two tokens of '*gewessen*' (see example 10), and one each of '*weßen*' (see example 11) and '*unsser*' (see example 12). There are also two tokens of '*Haußе*' (see example 13) in a letter by BOC and one token of '*Abwessenheit*' (see example 14) in a letter by WP. Neither KH's nor WP's letters contain examples of the corresponding SG form. However, two tokens of SG *Hause* occur in BOC's letters. Contrary to the findings of Elmentaler & Rosenberg (2015, p. 331), unvoiced /s/ does not appear in the ÖA corpus in word-initial position. In written SG, the graphemes <ss> and <ß> do not occur in word-initial position. Hence, the apparent absence of word-initial unvoiced /s/ may simply be due to orthographic conventions, i.e. word-initial unvoiced /s/ may well have

been a feature of the spoken German, but not of the written German of North Frisia. Unvoiced /s/ is not attested in any position by Elspaß (2005) or Langer (2013).

(9) KH: *Hausse* /haʊsə/
SG: *Hause* /haʊzə/
'home'

(10) KH: *gewessen* /gəveːsn/
SG: *gewesen* /gəveːzn/
'been'

(11) KH: *weßen* /veːsn/
SG: *wesen* /veːzn/
'being'

(12) KH: *unsser* /ʊnsɐ/
SG: *unser* /ʊnzɐ/
'our'

(13) BOC: *Hauße* /haʊsə/
SG: *Hause* /haʊzə/
'home'

(14) WP: *Abwessenheit* /apveːsnhaɪt/
SG: *Abwesenheit* /apveːznhaɪt/
'absence'

### *4.1.5 Caffe/Kaffee*

The German lexeme *Kaffee* ('coffee'), a loanword from French (*café*), has two standard realisations, /'kafe/ and /ka'feː/, and one diatopically-marked (northern) non-standard realisation, namely /'kafə/, i.e. first-syllable stress and word-final schwa in place of a full vowel (Elmentaler & Rosenberg, 2015, p. 179).[1] Today, this latter variant is found almost exclusively in the north and north-eastern regions of northern Germany, including North Frisia (Elmentaler & Rosenberg, 2015, p. 183). Langer (2013, p. 82) identifies a number of examples of the diatopically-marked variant in his corpus of soldiers' letters, in which the spellings '*Kaffe*' and '*Caffe*' appear in place of SG *Kaffee*. The ÖA corpus contains one example of the form '*Caffe*' in a letter by KH (see example 15). This spelling strongly suggests the non-standard northern pronunciation /'kafə/. SG *Kaffee* does not appear anywhere in the corpus.

[1] In northern German, /ka'feː/ refers to 'café', whereas /'kafe/ is 'coffee'.

(15) KH: *Caffe* /'kafə/
SG: *Kaffee* /'kafe/ or /ka'fe:/
'coffee'

## 4.2 Morpho-syntax

### *4.2.1 Case Morphology*

Examples of non-standard case morphology occur in the letters of all six writers and have a frequency of 13% in the ÖA corpus as a whole, i.e. 87% of instances case morphology appear in the standard form. While in all the letters the number of examples of standard case morphology outweighs the number of examples of non-standard case morphology, their relative proportions vary between writers. KH's letters contain the highest frequency of non-standard case morphology (30%), while BOC's contain the lowest (8%). Examples of non-standard case morphology are found in nouns (see example 16), pronouns (see example 17) and articles (see example 18). The majority of these examples concern the use of the accusative and dative cases. Usually, as in examples 16 and 18, the accusative is used in place of the dative, but in some cases, as in example 17, the dative is used in place of the accusative. This pattern broadly conforms to the findings presented by Langer (2013, p. 84). As Langer (2013, p. 85) notes, these 'uncertainties' (cf. Denkler & Elspaß, 2007, p. 97) with respect to case morphology may result from Low German (LG) interference, given that LG,
unlike SG, has only two cases, nominative and not-nominative, and therefore does not distinguish between two objective cases, i.e. accusative and dative (Lindow, Möhn, Niebaum, Stellmacher, Taubken & Wirrer, 1998, p. 144). Equally, the 'uncertainties' may result from North Frisian (NF) interference, given that NF also does not distinguish between two objective cases (Walker & Wilts, 2001, p. 289).

| | | | | |
|---|---|---|---|---|
| (16) KH: | *vor* | *14* | *tage*$_{\text{ACC.PL.}}$ | |
| | ago | 14 | days | |
| SG: | *vor* | *14* | *Tagen*$_{\text{DAT.PL.}}$ | |
| | ago | 14 | days | |
| '14 days ago' | | | | |

| | | | | | |
|---|---|---|---|---|---|
| (17) OJB: | *ihm*$_{DAT.}$ | *kenne* | *ich* | *nicht* |
| | him | know | I | not |
| SG: | *ihn*$_{ACC.}$ | *kenne* | *ich* | *nicht* |
| | him | know | I | not |
| 'I do not know him' | | | | |

(18) WP: *mit* *den*$_{\text{ACC.}}$ *Steuermann*
with the helmsman
SG: *mit* *dem*$_{\text{DAT.}}$ *Steuermann*
with the helmsman
'with the helmsman'

*4.2.2 Dative Nominal Possessive Constructions*

Dative nominal possessive constructions are a non-standard feature of a number of regional German varieties, including northern High German. In its usual form (see example 19), this construction comprises a dative nominal phrase and a possessive pronominal phrase (Zifonun, 2003, p. 98).

(19) *dem*$_{DAT.}$ *Mann* *sein*$_{POSS.PRON.}$ *Haus*
the man his house
'the man's house'

The ÖA corpus contains five examples of non-standard nominal possessive constructions, which appear in the letters of three different authors (KH, KW, BOC). Unlike the form outlined above, all five examples (see examples 20 to 24) feature a proper noun in place of the dative nominal phrase. This particular form of the nominal possessive construction is a diatopically-marked variant found in northern German (Denkler & Elspaß, 2007, p. 98). Similar examples are also attested by Elspaß (2005, p. 327) and Langer (2013, p. 85). The letters of KH, KW and BOC also contain nine examples of SG nominal possessive constructions (e.g. KW: *Riedels Frau*).

(20) KH: *Kresche* *ihr* *geld*
Kresche her money
SG: *Kresches* *Geld*
Kresche's money
'Kresche's money'

(21) KH: *Boÿ* *sein* *Frau*
Boy his wife
SG: *Boys* *Frau*
Boy's wife
'Boy's wife'

(22) KH: *Henrich* *sein* *Frau*
Henrich his wife
SG: *Henrichs* *Frau*
Henrich's wife
'Henrich's wife'

(23) KW: *Knudt* *seyn* *Brief*
Knudt his letter
SG: *Knudts* *Brief*
Knudt's letter
'Knudt's letter'

(24) BOC: *Carsten* *seine* *Aufführung*
Carsten his behaviour
SG: *Carstens* *Aufführung*
Carsten's behaviour
'Carsten's behaviour'

*4.2.2 Split Pronominal Adverbs*

Split pronominal adverbs are a characteristic non-standard trait of northern High German sentence structure (Denkler & Elspaß, 2007, p. 99). These constructions involve the separation of the pronominal and prepositional elements (see example 25) to form an adverbial bracket (Eroms, 2000, p. 136).

(25) *von* > *da…* *von*
thereof > there… of
'thereof'

The ÖA corpus contains two examples of split pronominal adverbs, both of which occur in a letter by KH (see examples 26 and 27). This variant is also attested by Langer (2013: 86). Split pronominal adverbs are a common feature of LG sentence structure (Lindow et al., 1998, p. 281). Similarly, they are also a feature of NF syntax (Hoekstra, 2001, p. 782). Hence, their appearance in the ÖA corpus may be the result of either LG- or NF-interference. The ÖA corpus contains 27 examples of standard, i.e. non-split, pronominal adverbs, three of which occur in the letters of KH.

(26) KH: *da* *wird* *wohl* *nicht* *mehr* *von*
there will surely not more of
SG: *es* *wird* *wohl* *nicht* *mehr* *davon* *geben*
there will surely not more thereof be
'there will surely not be any more of that'

(27) KH: *da* *war* *kein* *Brief* *beÿ*
there was no letter with
SG: *es* *gab* *kein* *Brief* *dabei*
there was no letter therewith
'there was no letter therewith'

*4.2.4 Wie/als and wie/wo*

SG *wie* can have either an interrogative (= 'how') or comparative (= 'as', see example 28) function. SG comparative constructions are also formed with *als* ('than', see example 29).

(28)

| *Er* | *ist* | *so* | *alt* | *wie* | *ich* |
|---|---|---|---|---|---|
| He | is | as | old | as | me |

'He is as old as me'

(29)

| *Er* | *ist* | *älter* | *als* | *ich* |
|---|---|---|---|---|
| He | is | older | than | me |

'He is older than me'

LG has no equivalent of SG *wie*; instead, the interrogative function is performed by *wo* and the comparative function by *as* (Lindow et al., 1998, pp. 216, 300). For this reason, native LG-speakers may use the phonologically-similar SG *als* where LG *as* would be used but SG *wie* would not, and vice-versa (Langer, 2013, p. 87). Similarly, native LG-speakers may use SG *wo* ('where') in place of SG *wie* in places where the phonologically-identical *wo* would be used in LG. The ÖA corpus contains examples of both types of LG interference. WP's letters contain three examples in which *wie* is used in place of SG *als* (see examples 30 to 32). These can be understood as instances of hypercorrection. WP's letters also contain two corresponding examples of SG *als* used 'correctly' in reporting past events. Non-standard use of SG *als*, analogous to LG *as*, does not appear in the ÖA corpus. The letter by HS contains a single example of LG *wo* used in place of the SG interrogative *wie* (see example 33). HS's letter contains no standard usage of interrogative *wie*.

(30)

| WP: | *wie* | *ich* | *mit* | *Riedel* | *war* |
|---|---|---|---|---|---|
| | how | I | with | Riedel | was |
| SG: | *als* | *ich* | *mit* | *Riedel* | *war* |
| | when | I | with | Riedel | was |

'when I was with Riedel'

(31)

| WP: | *wie* | *er* | *näher* | *hinzu* | *kam* |
|---|---|---|---|---|---|
| | how | he | nearer | to | came |
| SG: | *als* | *er* | *näher* | *hinzu* | *kam* |
| | when | he | nearer | to | came |

'when he came nearer'

(32)

| WP: | *wie* | *ich* | *an* | *Bord* | *kam* |
|---|---|---|---|---|---|
| | how | I | on | board | came |
| SG: | *als* | *ich* | *an* | *Bord* | *kam* |
| | when | I | on | board | came |

'when I came on board'

| | | | | | |
|---|---|---|---|---|---|
| (33) | HS: | *wo* | *weidt* | *das* | *herkömt* |
| | | where | far | that | comes |
| | SG: | *wie* | *weit* | *das* | *kommt* |
| | | how | far | that | comes |

'how far that comes'

## 5 Conclusion

The foregoing analysis demonstrates that the letters in the ÖA corpus exhibit a number of classic northern German linguistic features. In the realm of phonology these include *g*-spirantisation, *t*-apocope, affricate reduction and voiceless /s/. Non-standard case morphology and split pronominal adverbs are among the northern German morpho-syntactic features attested. The majority of these non- standard, diatopically-marked variants have been attested in other nineteenth- century northern German corpora and have also been shown to still exist today. Hence, the results suggest that the High German written (and presumably spoken) in nineteenth-century North Frisia shared a number of common linguistic traits with the High German written (and spoken) elsewhere in northern Germany. Several of these traits, such as non-standard case morphology, appear to have resulted from interference from Low German or North Frisian. Others may have arisen as a result of contact with other northern varieties of High German. While the range of non-standard diatopically-marked variants attested is reasonably wide, the number of examples of each variant is generally relatively small. This is presumably due to the fact that the authors would have had a high level of exposure to written standard High German through schoolbooks and religious texts and would thus have had a fairly clear idea of what they were aiming to recreate in their own writing. This is further supported by the presence, in many cases, of the co-occurring standard variants.

**Acknowledgements**

The research for this paper received funding from the Bristol Institute for Research in the Humanities and Arts (BIRTHA), University of Bristol (UK) and the School of Modern Languages, University of Bristol (UK). I am very grateful to the Ferring Stiftung (Alkersum auf Föhr, Germany), in particular Volkert F. Faltings and Reinhard Jannen, for their generous hospitality and intellectual insights. I also gratefully acknowledge the comments and suggestions made by Nils Langer and two anonymous reviewers on an earlier draft of this paper. Any errors and misjudgements that remain are, of course, solely my responsibility.

## References


Clyne, M. (1992). *Pluricentric languages: Differing norms in different nations*. Berlin, Germany: de Gruyter.

Denkler, M. & Elspaß, S. (2007). Nähesprachlichkeit und Regionalsprachlichkeit in historischer Perspektive. *Niederdeutsches Jahrbuch, 130*, 79-108

Elmentaler, M. & Rosenberg, P. (2015). *Norddeutscher Sprachatlas (NOSA). Band 1*. Hildesheim, Germany: Olms.

Elspaß, S. (2005). *Sprachgeschichte von unten: Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. Tübingen, Germany: Niemeyer.

Eroms, H.-W. (2000). *Syntax der deutschen Sprache*. Berlin, Germany: de Gruyter.

Faltings, V. F. (1992). Die mittelniederdeutsche Schreibtradition Nordfrieslands als Quelle der ältesten nordfriesischen Sprachüberlieferung. *Friesische Studien*, *1*, 53-98.

Hoekstra, J. F. (2001). Comparative aspects of Frisian morphology and syntax. In H. H. Munske (Ed.), *Handbuch des Friesischen/Handbook of Frisian Studies*, (pp. 775-785). Tübingen, Germany: Niemeyer.

Jensen, J. (1961). *Nordfriesland in den geistigen und politischen Strömungen des 19. Jahrhunderts (1797-1864)*. Neumünster, Germany: Wachholtz.

Koch, P. & Oesterreicher, W. (1990). *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*. Tübingen, Germany: Niemeyer.

Kohl, J. G. (1846). *Die Marschen und Inseln der Herzogthümer Schleswig und Holstein*. Dresden, Germany: Arnold.

Labov, W. (1994). *Principles of linguistic change. Volume 1: Internal factors*. Oxford, UK: Blackwell.

Langer, N. (2013). Norddeutsches in holsteinischen Soldatenbriefen von 1848-1850. *Niederdeutsches Jahrbuch, 136*, 73-96.

Langer, N. (2014). Language policies in the Duchy of Schleswig under Denmark and Prussia. In R. Healy & E. Dal Lago (Eds.), *The shadow of colonialism on Europe's modern past* (pp. 73-91). Basingstoke, UK: Palgrave Macmillan.

Lindow, W., Möhn, D., Niebaum, H., Stellmacher, D., Taubken, H. & Wirrer, J. (1998). *Niederdeutsche Grammatik*. Leer, Germany: Schuster.

Milroy, J. (2012). Sociolinguistics and ideologies in language history. In J. M. Hernádez-Campoy & J. C. Conde-Silvestre (Eds.), *The Handbook of Historical Sociolinguistics* (pp. 571-584). Oxford, UK: Blackwell.

Nobels, J. & van der Wal, M. (2012). Linking words to writers: Building a reliable corpus for historical sociolinguistic research. In N. Langer, S. Davies & W. Vandenbussche (Eds.), *Language and history, linguistics and historiography: Interdisciplinary approaches*, (pp. 343-362). Bern, Switzerland: Lang.

Rutten, G. & van der Wal, M. (2014). *Letters as loot: A sociolinguistic approach to seventeenth- and eighteenth-century Dutch*. Amsterdam, The Netherlands: Benjamins.

Walker, A. G. H. & Wilts, O. (2001). Die nordfriesischen Mundarten. In H. H. Munske (Ed.), *Handbuch des Friesischen/Handbook of Frisian studies*, (pp. 284-304). Tübingen, Germany: Niemeyer.

Zifonun, G. (2003). *Dem Vater sein Hut*. Der Charme des Substandards und wie wir ihm gerecht werden. *Deutsche Sprache*, *2*(3), 97-126.

# Arguments that Japanese [Cj]s are complex onsets: durations of Japanese [Cj]s and Russian [C$^{j}$]s and blocking of Japanese vowel devoicing

**Akitsugu Nogita**
University of Victoria
*akitsugu@uvic.ca*

This study examines whether the Japanese consonant-glide sequence [Cj]s as in [mjakɯ] 'pulse' are complex onsets /CC/ or palatalized consonants /C$^{j}$/ on the basis of duration. I compared the controversial Japanese consonant-cluster [Cj] analysis and uncontroversial Russian palatalized /C$^{j}$/ analysis in duration. The results indicate that Japanese [CjV]s are significantly longer than their [CV] counterparts, whereas Russian [C$^{j}$V]s are not significantly longer than their [CV] counterparts. Thus, in terms of duration, Japanese [Cj]s resemble consonant clusters /CC/ or /Cj/, and not Russian /C$^{j}$V/s. On the other hand, no arguments seem to exist to support that Japanese [Cj]s are palatalized single consonants based on the results. In addition to duration differences, I also assumed that if [Cj] is a consonant cluster, [j] in [CjɯC] would block [ɯ] devoicing/deletion. The results indicate that [j] blocks [ɯ] devoicing to some extent, but this may also be because infrequent morae may be less frequently devoiced. Therefore, the devoicing pattern does not support the Japanese complex onset hypothesis as strongly as the duration patterns.
*Keywords: Japanese consonant-glide sequence; Russian palatalized consonant; complex onset; high vowel devoicing*

## Introduction

There is a long-standing debate on whether [Cj] sequences, as in [ɾja.kɯ][1] 'omission' in Standard Japanese (SJ), are complex onsets /CC/ (/Cj/) or single palatalized consonants /C$^{j}$/. In this paper, I attempt to settle this long-standing debate on the basis of phonetic duration. In the first part of this paper, I examine 1) whether SJ [CjV] sequences are phonetically longer than their [CV] counterparts, 2) and if they are, whether the durational difference between SJ [CjV] morae and the [CV] counterparts is more similar to the durational

[1] Although the vowel [ɯ] in [ɾja.kɯ] in Standard Japanese is typically treated as an unrounded back vowel [ɯ], Nogita, Yamane, and Bird's (2013) ultrasound study revealed that this vowel pronounced in careful speech by linguistically naïve native SJ speakers is central or rather front with lip rounding and can be with unambiguous lip protrusion, so its actual realizations are [ɯ - ʏ]. I follow their recommended symbol [ɯ] or /ɯ/.

difference between Russian [CjV] syllables and the [CV] counterparts than to the difference between Russian [C$^{j}$V] syllables and the [CV] counterparts. I consider that a cross-linguistic comparison would shed light on this debate, especially with a language like Russian that has both uncontroversial palatalized /C$^{j}$/s, as in [s$^{j}$est$^{j}$] 'sit down' as well as uncontroversial consonant clusters /Cj/s, as in [sjest$^{j}$] 'eat up' (examples are from Kenstowicz, 1994, p. 42).

In the second part of this paper, I examine whether [j] in [Cj̥C] (C = voiceless consonant) blocks [u̥] devoicing in SJ. This additional experiment is designed based on feedback I received at a conference. SJ has a common rule that /u̥/ between voiceless consonants are regularly devoiced; i.e. /Cu̥C/ → [Cu̥C]. However, if [j] in [Cj] is found to block /u̥/ devoicing, I can argue that /u̥/ in /Cj̥u̥C/ is not between voiceless consonants and that [j] is an individual (voiced) consonant. I will examine this with nonsense words.

I am aware that there is also another argument that [j] is part of a complex nucleus /iV/ (/ / = non-syllabic/moraic), as some linguists (e.g., Hashimoto, 1984) propose. However, the goal of this study is only to reject the /C$^{j}$/ hypothesis, so I will not discuss the complex nucleus analysis.

## 1 Duration of Japanese [Cj] sequences

### 1.1 Brief introduction to Standard Japanese phonology

#### *1.1.1 The Japanese onset phoneme inventory*

This paper adopts Vance's (2008) and Larson-Hall's (2004) 19 Standard Japanese (SJ) onset inventory, as shown in Table 1, in which the so-called loanword sounds /ɕ, tɕ, ts, dz, ɸ/ are included.[2]

| | bilabial | dental | alveolar | alveopalatal | palatal | velar | glottal |
|---|---|---|---|---|---|---|---|
| plosive | p | t | | | | k | |
| | b | d | | | | g | |
| affricate | | ts | | tɕ | | | |
| | | dz | | dʑ | | | |
| fricative | ɸ | s | | ɕ | | | h |
| nasal | m | n | | | | | |
| liquid | | | ɾ | | | | |
| glide | w | | | | j | | |

*Table 1 SJ 19-onset-phoneme inventory*

*Note:* I excluded the consonants that occur only in coda positions or in the special/dependent mora positions, as they are irrelevant to my analysis.

[2] Grenon (2005) states that /dz, dʑ/, rather than /z, ʑ/, are underlying representations (UR) in SJ from a phonetic point of view. In this system, all the fricatives lack voicing contrasts while all the stops and affricates have them.

### *1.1.2 The Japanese palatal consonants [ɕ, tɕ, dʑ]*

How the so-called SJ palatal(ized) consonant series are treated and whether the loanword stratum is included have considerably varied in the analyses of the SJ consonant inventory. Since different assumptions of the onset inventory could lead to different conclusions regarding [Cj]s, deciding on an inventory is crucial.

Based on the [Cj]=/Cj/ assumption, Vance (2008) states that 11 of the 19 consonants can precede /j/, namely, /pj, bj, ɸj, mj, tj, dj, nj, ɾj, kj, gj, hj/, whereas sibilants and glides cannot, as sequences like */sj, ʦj, ɕj, tɕj, dʑj, dʒj, wj, jj/ are not attested. Like Vance (2008) and Larson-Hall (2004), I regard the so-called palatal(ized) series [ɕ, tɕ, dʑ] as single phonemes /ɕ, tɕ, dʑ/ as opposed to the other 11 /pj, bj, ɸj, mj, tj, dj, nj, ɾj, kj, gj, hj/ for two reasons related to articulation and phonotactics.

First, Hall (2000) proposes two types of palatalization: *'true' palatalization* which adds [-back] (tongue fronting), and *nonanteriorization* which adds [-anterior, +distributed] (shift from alveolar to postalveolar). For example, (1) and (2) show how /t, d, s, z, n, l/ change based on these two types of palatalization.

(1) /t, d, s, z, n, l/ → [$t^j$, $d^j$, $s^j$, $z^j$, $n^j$, $l^j$] by true palatalization

(2) /t, d, s, z, n, l/ → [ʧ, ʤ, ʃ, ʒ, ɲ, ʎ] by nonanteriorization

Since SJ [ɕ, tɕ, dʑ], as in [ɕakai] 'society', [otɕa] 'green tea', and [dʑama] 'interruption' are not realized as *[$s^j$, $t^j$, $d^j$], SJ [ɕ, tɕ, dʑ] are better treated as consonants with *nonanteriorization*.

Second, in SJ, [ɕ, tɕ, dʑ] contrast with the non-palatal counterparts [s, t, d] when preceding the vowel [i] (i.e. [ɕi, tɕi, dʑi] vs. [si, ti, di]). Other non-palatal consonants like [p, m, n, k, h], however, never contrast with [pj, mj, nj, kj, hj] before /i/, nor does *[ji] contrast with [i] (i.e. for example, [pi, ni, ki, i] vs. *[pji, nji, kji, ji]). In more details, Matsuzaki (1993) states that the /ti, di/ versus /tɕi, dʑi/ contrasts are unanimously accepted by linguists. In fact, there are (near) minimal pairs like /ti:/ ティー 'tea' versus /tɕii/ 地位 'status' as well as /diNkɯsɯ/ ディンクス 'double income no kids' versus /dʑiNkɯsɯ/ ジンクス 'jinx'. According to Nogita (2010), /si/ and /ɕi/ are marginally contrastive, as in /midʑɯhaɕi paɾɯsi(:)/ 水橋パルスィ (a name of a game character). There are also /ʦi, dzi/ as opposed to /tɕi, dʑi/, as in /eɾitsiN/ エリツィン 'Yeltsin' (a Russian name ) (Vance, 2008, p. 84) and /dzi:dzi:/ ズィー・ズィー (a name of a manga character).

It should also be noted that, [ɕ, tɕ, dʑ] before another front vowel [e] (i.e. [ɕe, tɕe, dʑe]) exist stably in loanwords, /tɕesɯ/ チェス 'chess', for example, whereas that of /je/ is not as stable (Takayama, 2005); /je/ is often replaced with /ie/, as in /ieɾo:/ イエロー 'yellow' rather than /jeɾo:/. Moreover, *[Cje]s, such as *[pje, kje], are not allowed (although some linguists acknowledge the existence of [hje, nje] (Matsuzaki, 1993)). All this suggests that [ɕe, tɕe, dʑe] behave differently from [je] and [Cje].

For these reasons, [ɕ, tɕ, dʑ] are phonologically treated as /ɕ, tɕ, dʑ/, but not as /sj, tj, dj (dʒj)/ in this paper. This means that the so-called *yōon* 拗音 (morae written with one regular-sized *kana* letter along with another small kana letter, such as <きゃ> (/kja/) or <しゃ> (/ɕa/)) in the regular Japanese term is not necessarily equivalent to /CjV/s. In other words, although [ɕV, tɕV, dʑV]s are *yōon* morae, I treat them as CV morae in this paper. Furthermore, In §2.3.4, I will provide further evidence that at least [dʑV]s behave like CV morae in terms of duration, as opposed to [CjV]s.

### 1.2 Previous studies of durations in SJ [CjV]

Durations in SJ [CjV]s have been discussed in previous studies. For example, Kida (1998) reports that phonetically [kja] is longer than [ka], which in turn is longer than [a]. Likewise, Yamaoka (2008) finds that SJ [j] in [ja] is longer than [j] in [bja, pja, mja], which in turn is longer than [j] in [kja, gja]. Note that these two authors do not provide acoustic data in their studies.

According to Parker (2012), one criterion distinguishing complex onsets and palatalized consonants is as follows: if [Cj] is a complex onset, duration of [j] is approximately as long as that of a single unambiguous onset glide. Based on this criterion, one might be inclined to the palatalized consonant hypothesis /C$^{j}$/ given that SJ [j] in [CjV] is shorter than [j] in [jV]. However, SJ is a mora-timed language, in which each mora is isochronous in SJ speakers' psychological timing units (but acoustically not necessarily isochronous) (Vance, 2008). In fact, according to Kohno's (1998) experiment of sentence reading, average *mora* duration in SJ is 145ms (S.D.=27.8ms, with S.D. meaning 'standard deviation'), while that of English and Spanish *syllables* are 244ms (S.D.=85.7) and 201ms (S.D.=74.0) respectively. The small S.D. values of the Japanese *morae* imply that the duration of SJ morae is relatively consistent at least in reading tasks compared to *syllables* in English and Spanish. This suggests that speakers may compress [Cj] in order to maintain the mora-timed rhythm, therefore resulting in a shorter durations of [j]. Another possible explanation can be that SJ has the [CjV]-[Ci.V] ('.' indicates a mora boundary) contrast, as in the minimal pair [kjo:] 'Buddhist sutra' versus [ki.o:] 'past illness' and [ki.jo:] 'appointment (to a position)'. Speakers may shorten [j] in [CjV] in order to make a perceptually clear contrast from [Ci.V]. In either case, in my review of the literature, I did not find studies that focused on acoustically to what extent [CjV]s are longer than [CV]s in SJ.

### 1.3 Experiment 1: durations of SJ [CjV]s and the [CV] counterparts

My first question is whether SJ [CjV] morae duration is longer than that of the [CV] counterparts. To examine it, I instructed native SJ speaking participants to

pronounce these morae in nonsense words embedded in a sentence, and measured the duration[3].

### *1.3.1 Participants*

34 (11 male and 23 female) native SJ speakers, aging from 15 to 59, participated in the experiment. All the participants were reportedly born and raised in or near Tokyo: more specifically, Tokyo, Kanagawa, Saitama, Chiba, Yamanashi, and Gunma. As my experiment was conducted in Victoria, British Columbia, Canada, the participants were recruited in Victoria.

### *1.3.2 Stimuli and procedure*

The target morae were [gʲja, nja, mja, bja, ɾʲja] versus [ga, na, ma, ba, ɾa] respectively, and the control pair was [d͡ʑa] (palatal /CV/) versus [d͡za] (non-palatal /CV/). The durations of these morae were compared. Voiced consonants were chosen for the stimuli with the purpose of avoiding a possibility of vowel devoicing/deletion when adjacent to voiceless consonants, which might affect the results. To avoid homorganicity, these morae were embedded in a 4-mora nonsense compound word [_tá-bako] when the target consonant was non-coronal, for example, [m̲atá-bak̲o], and [_ká-bako] when the target consonant was coronal, for example, [n̲ak̲á-bako]. The participants were instructed to place the phonemic pitch accent (or a high tone) on the second mora in these four-mora nonsense words; [bako] is from a real word 'box', and in real compound words, the mora before [bako] is the default accented mora (e.g. [omot͡ɕá-bako] 'toy box'), so the place of the accented mora in [_tá-bako]/[_ká-bako] is natural to SJ speakers. These nonsense words were embedded in the sentence [watad͡ʑitat͡ɕiwa _ desɯ] 'We are _.'

These stimulus sentences were printed on a sheet with regular Japanese orthography (see Appendix for the stimulus sentences). The participants read aloud each sentence three times in natural speed and only the second token was analyzed. They were encouraged to practice all the nonsense words once beforehand to familiarize themselves with the words, which also allowed me to correct their obvious mistakes if any. Recording was done in the soundproof booth in the UVic Phonetics Lab with a microphone, SONY ECM-MS908C, and the Software Audacity set at 44100Hz and 32-bit float.

[3] In my analysis, I use the term 'mora' rather than 'syllable' since the necessity of the syllable nodes in the prosodic hierarchy in SJ is debatable (e.g., Labrune, 2012). The difference between morae and syllables is that a mora counts a coda consonant and the second half of a long vowel/diphthong; as an example, the /CCV:C/ syllable has three morae, /CCV/ + /:/ + /C/.

### *1.3.3 Data analysis*

Since it is difficult to spot the boundary between [j] and the following vowel, the durations of the whole morae were measured. In addition to the absolute duration, the ratio between the target mora and the following mora (e.g., the ration between [gja] and [tá] in [gjatábako]) was compared in consideration of the speech rate. For morae with stop/affricate/flap onsets, the measurement points were from the beginning (zero crossing) of the closure to the end of the last pitch pulse. For nasal onsets, the beginning points were the beginning of the first pitch pulse that shows a sudden change of the waveform pattern (see Figure 1).

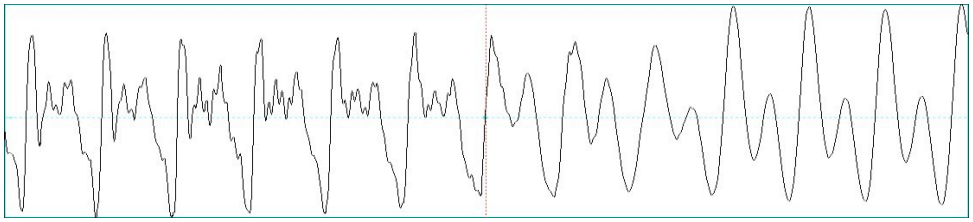


*Figure 1. The measurement point of the beginning of a nasal consonant; the vertical red dot line indicates the measurement point.*

The analysis was done using the phonetic software PRAAT. All the data were measured twice, and the consistency across measurement 1 and 2 was 86.5%. Despite the relatively low consistency, [Cja]s were longer than [Ca]s by 12ms to 25ms on average, as shown in Table 2 in §2.3.4, so even if measurement points in some tokens were wrong by one pitch pulse (roughly 3ms to 7ms depending on the gender), the overall results would still be the same.

### *1.3.4 Results and discussion*

Table 2 shows the average durations and average ratios.

| Mora | dza | dʑa | ga | gja | na | nja | ɾa | ɾja |
|---|---|---|---|---|---|---|---|---|
| Duration[a] | 147 | 144 | 130 | 142 | 134 | 156 | 115 | 132 |
| p-value[b] | p=0.28 | | *p=0.00062 | | *p=3E-08 | | *p=6E-05 | |
| Duration of tá/ká | 138 | 141 | 147 | 140 | 143 | 144 | 147 | 143 |
| p-value | p=0.39 | | *p=0.042 | | p=0.60 | | p=0.15 | |
| Ratio % | 109% | 104% | 89% | 103% | 95% | 109% | 79% | 93% |
| p-value | p=0.18 | | *p=4E-30 | | *p=1E-07 | | *p=2E-06 | |

| Mora | ma | mja | ba | bja |
|---|---|---|---|---|
| Duration[a] | 135 | 160 | 136 | 161 |
| p-value[b] | *p=1E-08 | | *p=6E-08 | |
| Duration of tá/ká | 142 | 138 | 140 | 138 |
| p-value | p=0.15 | | p=0.29 | |

*Table 2 Average durations of unaccented [C(j)a] and the following [tá/ká], and average ratios between [C(j)a] and the following [tá]/[ká] in Standard Japanese*
*Note*:[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [C(j)a]s and their [Ca] counterparts. Asterisk* indicates a significant difference.

The [gja, nja, mja, bja, ɾja] morae were consistently longer than corresponding [ga, na, ma, ba, ɾa]; note that according to the [C(j)a]-[tá/ká] ratios, in only 20 (out of 170) cases, [CV] was longer than its [CjV] counterpart of the same speaker (6 in [ga-gja], 5 in [na-nja], 2 in [ma-mja], 3 in [ba-bja], 4 in [ɾa-ɾja]). Overall, both absolute durations and the [C(j)a]-[tá/ká] ratios of [gja, nja, mja, bja, ɾja] were highly significantly longer than the [ga, na, ma, ba, ɾa] counterparts. In contrast, between [ʣa] and [ʥa], unexpectedly, the non-palatal [ʣa] was slightly longer than the palatal [ʥa], but the difference was not significant. Since the non-palatal versus palatal pair [ʣa]-[ʥa] did not show a difference in duration, at least in SJ, the generalization that palatal consonants are longer than the non-palatal counterparts was not observed. In other words, I can argue that the fact that [gja, nja, mja, bja, ɾja] are significantly longer than [ga, na, ma, ba, ɾa] is not due to the nature of palatal consonants, but because [gja, nja, mja, bja, ɾja] consist of three segments /CCV/ while [ga, na, ma, ba, ɾa] as well as [ʣa, ʥa] consist of two segments /CV/.

### *1.3.5 Other findings that support previous studies*

One result of this experiment is that [gja] was longer than [ga] only by 12ms, while [mja, bja] were longer than [ma, ba] by 25ms. This is consistent with Yamaoka's (2008) statement that [j] in [pj, bj, mj] is longer than [j] in [kj, gj]. The occasions that [mja, bja] failed to be longer than [ma, ba] were two and three times respectively, while the occasions that [gja] failed to be longer than [ga] were as many as six times. In other words, [mja, bja] showed long durations more consistently than [gja] did.

I suspected that there may be compensatory shortening in the following mora [tá/ká] to maintain the mora-timed rhythm, that is, if a [CjV] mora is long, speakers might shorten the following mora to keep the same duration of the whole word. According to Table 2, except for [ká] after [nja, na], the mora [tá/ká] after [Cja] was slightly shorter than [tá/ká] after [Ca]. However, a statistically significant shortening was found only in [tá] after [ga, gja]. Moreover, the difference between [tá] after [gja] and [tá] after [ga] was only 7ms on average, which may not be perceptually salient. Therefore, there was no solid evidence for compensatory shortening. Based on my perceptual impression, the nonsense words containing [Cja] sometimes sounded slower than those without [Cja]. Given this impression, I also suspected a global slowdown by lengthening the following mora in order not to make the [CjV] mora sound deviant from the surrounding morae in duration as an alternative strategy for the mora-timed rhythm. However, as Table 2 shows, this was not the case either. That is, neither compensatory shortening of the following mora nor a global slowdown was observed in this experiment. These results agree with Warner and Arai (2001), who do not find any compensation related to the mora-timed rhythm either. This means that my perceptual impression that the nonsense words containing [Cja] sounded slower than those without [Cja] is only because of the longer duration of [Cja]. Still, Warner and Arai (2001, p. 1149) find a high correlation between duration and the number of morae in spontaneous speech by 11 native Japanese (not necessarily SJ) speakers (the r-value ranging from 0.701 to 0.931). Thus, there is a possibility that short duration in [j] in [CjV] comes from an attempt to keep the mora-timed rhythm.

### 1.4 Intermediate conclusion

To the question of whether SJ [CjV] morae are phonetically longer than their [CV] counterparts, the answer is yes. [CjV]s are highly significantly longer than the [CV] counterparts. Since the palatal [dʑ] was not longer than the non-palatal counterpart [dz], the reason that [CjV]s are longer than corresponding [CV] is not because of the long duration of palatal consonants. Instead, this suggests a possibility that [CjV]s consist of three segments.

## 2 Duration of Russian [$C^j$] and [$C^{(ɣ)}$]

### 2.1 The second research question

In the last section, I concluded that palatal consonants are not phonetically longer than their non-palatal counterparts. However, only the coronal affricate [dʑ] and [dz] were examined. In this section, I examine whether consonants with true palatalization (see §2.1), such as [$n^j$, $m^j$, $b^j$, $r^j$], do not show longer durations than corresponding [n, m, b, r]. If [$n^j$, $m^j$, $b^j$, $r^j$] are clearly longer than [n, m, b, r] respectively, then it can be concluded that SJ [nj, mj, bj, ɾj] are longer than [n, m, b, ɾ] because of the longer duration of palatalized consonants. In other words, SJ [nj, mj, bj, ɾj] can be interpreted as single palatalized consonants /$n^j$, $m^j$, $b^j$, $ɾ^j$/. In contrast, if [$n^j$, $m^j$, $b^j$, $r^j$] are not longer than [n, m, b, r], it can be another piece of evidence that SJ [nj, mj, bj, ɾj] are consonant clusters /nj, mj, bj, ɾj/. In order to examine the duration of [$C^j$]s, I investigate Russian, which has unambiguous contrastive palatalized consonants /$C^j$/s (also called soft consonants, мягкие согласные) as opposed to plain/velarized consonants /$C^{(ɣ)}$/s (also called hard consonants, твердые согласные). In this paper, I omit the velarized symbol /$^ɣ$/ for hard consonants.

## 2.2 Brief introduction to Russian phonology

Russian has phonologically contrastive palatalized consonants and plain/velarized counterparts, as in [m$^{j}$at$^{j}$] 'to rumble', [mat] 'checkmate', [m$^{j}$at] 'rumpled', and [mat$^{j}$] 'mother' (Kenstowicz, 1994, p. 41). Table 3 shows the Russian consonant inventory from Padgett (2003b).

| | labial | | dental | | post-alveolar | | palatal | velar | |
|---|---|---|---|---|---|---|---|---|---|
| plosive | p | p$^{j}$ | t | t$^{j}$ | | | | k | k$^{j}$ |
| | b | b$^{j}$ | d | d$^{j}$ | | | | g | g$^{j}$ |
| fricative | f | f$^{j}$ | s | s$^{j}$ | ʃ | ʃ$^{j}$: | | x | x$^{j}$ |
| | v | v$^{j}$ | z | z$^{j}$ | ʒ | | | | |
| affricate | | | ʦ | | | ʧ$^{j}$ | | | |
| nasal | m | m$^{j}$ | n | n$^{j}$ | | | | | |
| liquid | | | l | l$^{j}$ | | | | | |
| | | | r | r$^{j}$ | | | | | |
| glide | | | | | | | j | | |

*Table 3 Russian consonant inventory from Padgett (2003b)*

Padgett (2003a) mentions that if a language has palatalized consonants, plain counterparts tend to be velarized for the sake of perceptually clear distinctiveness. For example, Kochetov (2002) observes that Russian [p$^{j}$] in [ap$^{j}$á] involves tongue body fronting and raising, while [p] in [apá] involves partial velarization (or pharyngealization), that is tongue body backing but without tongue body raising. Still, according to Litvin's (2014) ultrasound study, clear presence of velarization/pharyngealization in plain consonants is debatable. In any case, palatalized-plain/velarized contrasts are [-/+back] contrasts (Padgett, 2003a). I will examine durations of [-/+back] consonants.

In addition, Russian [C$^{j}$]s also phonologically contrast with consonant cluster [Cj]s as in [s$^{j}$est$^{j}$] 'sit down' versus [sjest$^{j}$] 'eat up' (Kenstowicz, 1994, p. 42). This contrast is what Japanese lacks. Besides, Russian also has a syllabicity contrast, such as [di.a] versus [dja] (Padgett, 2008), just as Japanese does.

## 2.3 Experiment 2: durations of Russian [C]s, [C$^{j}$]s, and [Cj]s in unstressed positions

### *2.3.1 Participants*

I recruited eight (3 male and 5 female) native Russian speakers, aging from 18 to 29, and reportedly born and raised in Russia (European Russia, the South-West area, and Perm), Kazakhstan, and Belarus. Again, all the participants were recruited in Victoria, British Columbia, Canada.

### *2.3.2 Stimuli, procedure, and analysis*

The three groups of target syllables were [na, ma, ba, ra], [nʲa, mʲa, bʲa, rʲa], and [nja, mja, bja, rja], which were the counterparts of the morae used in the Japanese experiment respectively. In the Russian experiment, the velar [gʲa] was discarded in analysis (although they were also recorded). This is because in Russian, while the hard-soft contrast in velars before /a, o, u/ is present in a fair number of loanwords, there is some dispute whether velars are contrastive in palatalization (Padgett, 2003b) and the speakers may not be familiar with it as well as the other palatalized consonants.

Like in the Japanese experiment, the stimulus syllables were embedded in the nonsense word template /_kábako/ or /_tábako/. Again, to avoid homorganicity, the template was /_tábako/ when the target consonant was non- coronal (e.g., /matábako/) and /_kábako/ when the target consonant was coronal (e.g., /nakábako/)[4]. The participants were instructed to place the phonemic stress on the second syllable in order to replicate the Japanese experiment[5]. These nonsense words were embedded in the sentence /mi _/ (or /mɨ _/, depending on the phonological interpretation) 'We are _.'

The stimuli were printed on a sheet with regular Cyrillic Russian orthography (see Appendix for the stimuli). The recording procedure and data analysis were done in the same way as in the Japanese experiment. The only difference is that in the Russian experiment, participants read aloud each sentence five times in natural speed and the middle three tokens were analyzed. This is for compensating my limited access to Russian speakers in Victoria.

### *2.3.3 Results and discussion*

Table 4 shows the average durations and average ratios regarding [CV]s and [CʲV]s in unstressed syllables. Table 5 shows those of [CV]s and [CjV]s.

[4] In the section of the Japanese experiment, I used the phonetic brackets [ ], as whether [Cj] is phonologically /Cʲ/ or /Cj/ had not been decided yet. In the Russian experiment, however, I use the phonemic brackets / /, since the status of the Russian consonants at issue is already uncontroversial. In addition, unstressed /a/ and /o/ are neutralized so that the orthographically indicated vowels in the stimuli can be different from the actual realizations. Therefore, using / / for the Russian stimulus sentences is more suitable.

[5] When pitch is involved, I use the term *accented*, and otherwise *stressed*. Since the Japanese pitch accent involves pitch, I call it *accented*. Since the Japanese pitch accent system can be categorized as a type of the tone system rather than the stress system (Hyman, 2006), the Russian experiment cannot exactly be a replication of the Japanese one in terms of stress/accent patterns.

| Syllable | na | nʲa | ma | mʲa | ba | bʲa | ra | rʲa |
|---|---|---|---|---|---|---|---|---|
| Duration[a] | 130 | 139 | 142 | 147 | 150 | 157 | 109 | 106 |
| p-value[b] | *p=0.012 | | p=0.147 | | p=0.228 | | p=0.450 | |
| Duration of tá/ká | 197 | 197 | 198 | 193 | 192 | 189 | 196 | 194 |
| p-value | p=0.943 | | p=0.420 | | p=0.620 | | p=0.608 | |
| Ratio % | 68% | 74% | 75% | 79% | 82% | 86% | 58% | 56% |
| p-value | *p=0.006 | | p=0.066 | | p=0.365 | | p=0.480 | |

*Table 4 Average durations of [C(ʲ)a] and the following [tá/ká], and average ratios between [C(ʲ)a] and the following [tá]/[ká] in Russian*

*Note*:[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [Ca]s and their [C(ʲ)a] counterparts. * indicates a significant difference at the $p<0.05$ level.

| Syllable | na | nja | ma | mja | ba | bja | ra | rja |
|---|---|---|---|---|---|---|---|---|
| Duration | 130 | 196 | 142 | 200 | 150 | 197 | 109 | 177 |
| p-value | *p=2E-08 | | *p=3E-10 | | *p=5E-06 | | *p=2E-11 | |
| Duration of tá/ká | 197 | 207 | 198 | 203 | 192 | 198 | 196 | 208 |
| p-value | p=0.058 | | p=0.31 | | p=0.43 | | p=0.10 | |
| Ratio % | 68% | 97% | 75% | 103% | 82% | 103% | 58% | 87% |
| p-value | *p=6E-10 | | *p=5E-07 | | *p=4E-06 | | *p=2E-09 | |

*Table 5 Average durations of [C(j)a] and the following [tá/ká], and average ratios between [C(j)a] and the following [tá]/[ká] in Russian*

According to the [Cʲ]-[C] comparison in Table 4, unlike the case of highly significant differences in all the Japanese [Cj]-[C] pairs, these Russian [Cʲ]-[C] pairs did not show significant differences in both absolute durations and the [C(ʲ)a]-[tá/ká] ratios, except that [nʲ] was significantly different from [n] in both the absolute duration ($p<0.05$) and in the ratio ($p<0.01$). It should be noted, however, that [nʲa] is longer than [na] only by 9 ms in the absolute duration and 6% in ratio on average, while the Japanese [nja, mja, bja, ɾja] were longer than corresponding [na, ma, ba, ɾa] by 17-25ms in duration and 14-21% in ratio on average[6]. Likewise, Russian [mʲ] was almost significantly longer than [ma] in the [mʲa]-[tá] ratio ($p=0.066$) (not significant in the absolute duration), but only by 4%. The differences in these two cases may not be perceptually salient. Therefore, there is no clear evidence indicating that palatalized consonants are longer than plain counterparts. Finally, I want to point out that while Japanese [CjV]s are impressionistically longer or slower than [CV]s, this is not the case with Russian [CʲV]s.

[6] Regarding the difference between Russian [nʲa] and [na] as opposed to the difference between Japanese [nja] and [na], one reviewer asked whether the difference in Japanese was statistically bigger than those in Russian. The two-tailed two-sample t-test showed that the difference in Japanese was significantly bigger than that in Russian ($p=0.013$ in duration and $p=0.005$ in ratio).

The question which deserves further discussion is why only [n$^{j}$a] was significantly longer than [na]. This may be because the preceding segment is the particular vowel /i/. Since the place of articulation of /i/ and that of /n$^{j}$/ is close, there may be partial assimilation in the end part of /i/, which ends up with slight lengthening of /n$^{j}$/. As a piece of evidence, while Figure 2 shows that the boundary between /i/ and /n/ is very clear, Figure 3 shows that the boundary between /i/ and /n$^{j}$/ by the same speaker is not as clear, and all the speakers show the same tendency. These data seem to show this assimilation in the /in$^{j}$/ sequence.

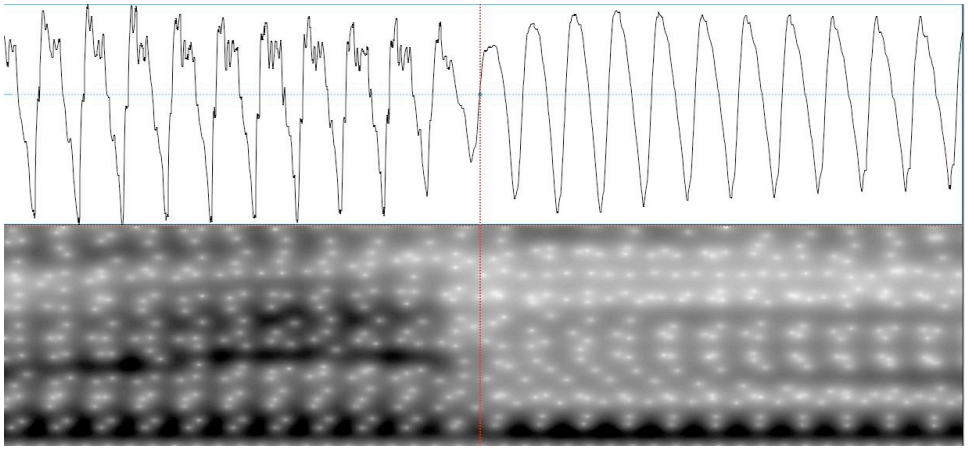



*Figure 2 Boundary between /i/ and /n/ in /mi nakábako/*

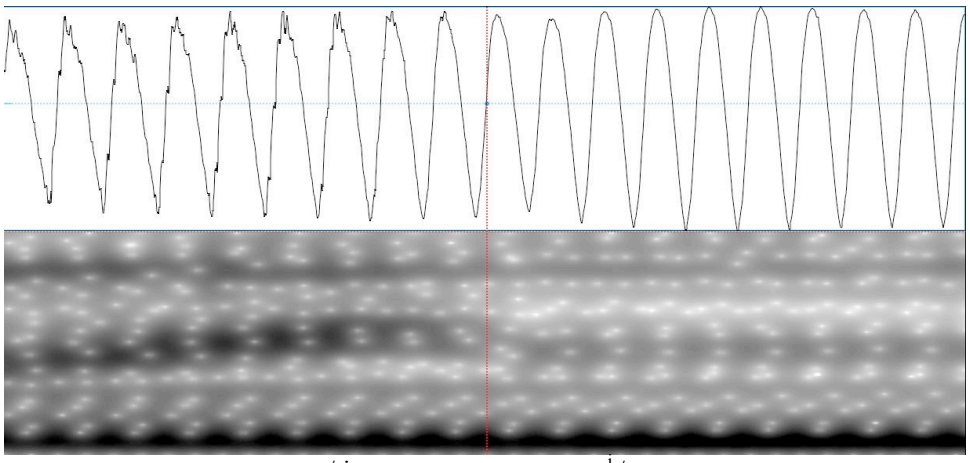



*Figure 3. Boundary between /i/ and /n$^{j}$/ in /mi n$^{j}$akábako/*

As for the [CjV]-[CV] comparison shown in Table 5, it is more straightforward. Table 5 indicates that [CjV]s are highly significantly longer than the [CV] counterparts. In the [C(j)a]-[tá/ká] ratios, only in one occasion ([mja- ma]) (out of the total 96) that [CjV] failed to be longer than the [CV] counterpart for the same speaker. In the absolute durations, only two occasions (both in [bja-

ba]) showed that [CjV] failed to be longer. The differences in the absolute durations between [CjV]s and [CV]s are 47-68ms and those of the [C(j)a]-[tá/ká] ratios are 21%-29%. Compared to the standard Japanese (SJ) cases (17-25 ms and 14-21%), Russian [CjV]s are longer than the [CV] counterparts more consistently and by larger degrees. This can be analyzed that the phonological contrasts between /CjV/s and the /C$^{j}$V/ counterparts need to be perceptually salient in Russian in order to maintain the contrasts. In contrast, SJ lacks the phonological /CjV/-/C$^{j}$V/ contrast, and therefore there is no motivation to lengthen [CjV]s. Moreover, SJ [CjV]s have to contrast with [Ci.V]s and have to keep the mora-timed rhythm. Thus, [CjV] with a long duration is rather not preferable. Therefore, the above analysis suggests that the fact that Japanese [CjV]s are longer than [CV]s less consistently and by smaller degrees than Russian counterparts cannot be the evidence that Japanese [Cj]s are phonologically /C$^{j}$/s.

As for the following [tá/ká] syllables, there were no significant differences between [tá/ká] following [CV]s and [tá/ká] following [C$^{j}$V]s, or between [tá/ká] following [CV]s and [tá/ká] following [CjV]s, suggesting that neither global slow down nor compensatory shortening is observed. This is the same as Japanese.

## 2.4 Experiment 3: durations of Russian [C]s, [C$^{j}$]s, and [Cj]s in stressed positions

### *2.4.1 Possible stress effect*

There is a possibility that the syllables in the experiment above were pronounced as unstressed syllables and their length differences between /C$^{j}$V/s and /CV/s may have been therefore reduced. To test this possibility, in this section, I examine Russian syllables in a stressed position.

### *2.4.2 Participants, stimuli, and procedure*

The participants in this experiment were the same as experiment 2. The target syllables were also [na, ma, ba, ra] versus [n$^{j}$a, m$^{j}$a, b$^{j}$a, r$^{j}$a] versus [nja, mja, bja, rja], except that they were stressed. In experiment 2, the stimuli used were replication of the Japanese nonsense words, which should have sounded very foreign for the Russian participants. In this experiment, the stimulus nonsense words were designed to be more Russian-like. The stressed target syllables were embedded in /__tap/ or /__kap/, for example, /n$^{j}$ákap/ and /m$^{j}$átap/. Again, /t/ and /k/ were alternated to avoid homorganicity. These stimulus words were embedded in /ona hot$^{j}$ela __ / 'She wanted __.' This time, the vowel immediately before the target syllable is /a/ instead of /i/ in order to avoid potential assimilation with the following palatalized consonant. The participants pronounced each sentence five times and the middle three were analyzed. Only the /ta/ or /ka/ was measured in the second syllable, and the coda /p/ was excluded.

*2.4.3 Results and discussion*

Table 6 shows the average durations and average ratios regarding [CV]s and [C$^{j}$V]s in stressed syllables. Table 7 shows those of [CV]s and [CjV]s.

| Syllable | na | n$^{j}$a | ma | m$^{j}$a | ba | b$^{j}$a | ra | r$^{j}$a |
|---|---|---|---|---|---|---|---|---|
| Duration[a] | 188 | 195 | 196 | 211 | 208 | 215 | 141 | 153 |
| p-value[b] | p=0.055 | | *p=0.0002 | | *p=0.025 | | *p=0.005 | |
| Duration of tá/ká | 143 | 147 | 145 | 149 | 148 | 148 | 149 | 151 |
| p-value | p=0.177 | | p=0.163 | | p=0.848 | | p=0.667 | |
| Ratio % | 136% | 138% | 142% | 150% | 148% | 155% | 98% | 104% |
| p-value | p=0.533 | | p=0.061 | | *p=0.022 | | p=0.106 | |

*Table 6 Average durations of stressed [C(j)á] and the following [ta/ka], and average ratios between [C(j)á] and the following [ta]/[ka] in Russian*

*Note.*[a] The durations are given in milliseconds. [b] p-values were calculated by the two-tailed paired t-test between [Ca]s and their [C(j)a] counterparts. * indicates a significant difference at the p<0.05 level.

| Syllable | na | n$^{j}$a | ma | m$^{j}$a | ba | b$^{j}$a | ra | r$^{j}$a |
|---|---|---|---|---|---|---|---|---|
| Duration[a] | 188 | 250 | 196 | 258 | 208 | 255 | 141 | 213 |
| p-value[b] | *p=4E-11 | | *p=1E-10 | | *p=1E-08 | | *p=7E-10 | |
| Duration of tá/ká | 143 | 157 | 145 | 147 | 148 | 148 | 149 | 153 |
| p-value | *p=0.007 | | p=0.554 | | p=0.859 | | p=0.289 | |
| Ratio % | 136% | 166% | 142% | 185% | 148% | 182% | 98% | 145% |
| p-value | *p=3E-05 | | *p=6E-08 | | *p=1E-05 | | *p=2E-08 | |

*Table 7 Average durations of stressed [C(j)á] and the following [ta/ka], and average ratios between [C(j)á] and the following [ta]/[ka] in Russian*

As shown in Table 6, interestingly, in stressed syllables, [C$^{j}$V]s are significantly longer than [CV]s in absolute durations, except that [n$^{j}$á] is only marginally significantly longer than [ná]. Moreover, the difference is 7ms to 15ms on average, so at least the [m$^{j}$á]-[má] and [r$^{j}$á]-[rá] differences might be perceptually noticeable (and according to my perceptual impression, [m$^{j}$á] sometimes sounded slightly longer than [má]). However, as mentioned above, absolute durations vary depending on the speech rate, so that ratios with the adjacent syllable would be more reliable. In fact, the [C$^{j}$á]-[ta/ka] ratios and the [Cá]-[ta/ka] ratios were not significantly different, except that [b$^{j}$á] was significantly longer than [bá] at the p<0.05 level. Also, the [m$^{j}$á]-[má] difference was marginally significant. However, the statistically significant difference between [b$^{j}$á] and [bá] was only 7%. Recall that the Japanese [nja, mja, bja, ɾja] were longer than [na, ma, ba, ɾa]

by 14-21% on average even in an unaccented low tone position with the mora-timed rhythm restriction[7]. This means that even the significant values in stressed syllables in Russian were much less obvious than the Japanese values. What is intriguing is that this time, the stressed [n$^{j}$á]-[ná] difference was not significant. This suggests that the significance in the unstressed [n$^{j}$a]-[na] difference in the previous experiment was likely because of partial assimilation of the preceding /i/ à [in$^{j}$]. In other words, the significance was only accidental. So I conclude that in Russian, only palatalized bilabial consonants [m$^{j}$] and [b$^{j}$] in stressed syllables tend to be slightly longer than the plain/velarized counterparts [m] and [b], but the difference is much less than the difference between Japanese [mj, bj] and [m, b].

Table 7 shows, unsurprisingly, that [CjV]s were significantly longer than the [CV] counterparts. In the [C(j)á]-[ta/ka] ratios, in five cases (1 in [mja], 2 in [bja], and 2 in [nja]) out of 96, [CV] was longer than the [CjV] counterpart by the same speaker, and in the absolute durations, only in one case (in [bja]) [CV] was longer. The differences in the absolute durations between [CjV]s and [CV]s were 47-72 ms and those of the [C(j)a]-[tá/ká] ratios were 30%-47%. These data indicate that in stressed positions the differences between [CjV]s and [CV]s become larger than in unstressed positions. Still, it is interesting that in Russian [CjV]s are occasionally shorter than [CV]s, suggesting that the fact that Japanese [CjV]s occasionally become shorter than [CV]s does not mean that Japanese [Cj]s are phonologically /C$^{j}$/s. Again, compensatory shortening of the following syllable or global slowdown is not observed, except that stimuli with [njá]s show global slow down.

### 2.5 Intermediate conclusion

My experiments indicate that while in both Japanese and Russian [CjV]s were highly significantly longer than the [CV] counterparts, Russian [C$^{j}$V] versus [CV]s did not show the same pattern. This suggests that Japanese [CjV]s are more similar to Russian /CjV/s than to Russian /C$^{j}$V/s. My experiments also show that Russian stressed bilabial [C$^{j}$]s may be slightly longer than their [C] counterparts, but the difference is not as significant as the difference between Japanese unaccented [Cj]s and [C]s. These phenomena seem to be better explained by suggesting that Japanese [CjV]s consist of three segments, rather than that longer durations are the nature of palatalized consonants. At least, there is no duration-related evidence that Japanese [Cj]s are closer to /C$^{j}$/s than they are to /Cj/s.

[7] Again, to answer one reviewer's question whether the difference between Japanese [bja] and [ba] (25ms, 20%) was statistically more significant than that of Russian [b$^{j}$a] and [ba] (7ms, 7%), I ran the t-test, which showed that the difference in Japanese was significantly bigger ($p<0.001$ for duration and $p=0.002$ for ratio).

## 3 Possible blocking of high vowel devoicing in Standard Japanese

### 3.1 Background and research question

When presenting the research on the Japanese [Cj] at the annual conference of The Phonological Society of Japan in 2013, I received feedback that if [j] blocks vowel devoicing, it can be evidence that [Cj] is a consonant cluster. Based on this feedback, I designed an additional experiment.

My research question is whether [j] in [Cj] blocks High Vowel Devoicing/Deletion (HVD) in SJ. As a common generalization, the short high vowels /ʉ/ and /i/ between voiceless consonants are typically devoiced/deleted in SJ, i.e. [ʉ] and [i] in /C̥ʉC̥, C̥iC̥/ (C=voiceless consonant) become [C̥ʉC̥, C̥iC̥] or [CC, CC]. If one of the consonants adjacent to [ʉ] or [i] is voiced, as a common generalization, HVD does not occur. Therefore, in [jʉC̥] contexts, /ʉ/ is not likely to be devoiced/deleted since [j] is a voiced consonant. Likewise, if [j] in [CjV] is the same consonant as [j] in [jV], the [CjʉC̥] contexts would block HVD of /ʉ/. Recall that since *[ji] and *[Cji] are disallowed in SJ (see §2.1), *[CjiC̥] does not occur, so I examine only [CjʉC̥].

In my review of the literature, only Kondo (2000) attempted to examine whether [Cj] is a complex onset or a palatalized consonant by observing HVD blocking in [CjʉC̥] contexts in real words, such as /okjʉpe:ɕoN/ ‘occupation’ (p. 135). Her finding shows that [j] in [CjʉC̥] blocked HVD 46 times out of 56 occasions in the word reading task by four native SJ speakers. If her results can be generalized, [Cj] would be an unambiguous complex onset because it can be explained that the voiced [j] changes the environment where [ʉ] is sandwiched between voiceless consonants. However, Kondo also mentioned that she used only real words, but [CjʉC̥] does not frequently occur in the actual vocabulary. Moreover, while an unaccented /ʉ/ in /C̥ʉS/ (S=voiceless stop) causes HVD nearly 100% of the time in SJ (Fujimoto & Kiritani, 2003), when /ʉ/ bears a phonemic accent nucleus (/C̥ʉC̥/) or when there are two consecutive devoiceable morae (/C̥ʉC̥ʉC̥/ etc.), it is not always devoiced (Kondo, 2000). In Kondo’s (2000) experiment, the words had environments where HVD does not always occur, such as /kjʉpʉra/ ‘cupra’ with an accented /ʉ/ (p. 135). In fact, Kondo (2012, in personal communication) said that her observation was preliminary and cannot be generalized. Furthermore, contrary to Kondo (2000), according to Shinohara (2012), in prescriptive Japanese, /Cjʉ(C)/ morae (/kjʉ, hjʉ, pjʉ/) are supposed to be devoiced in order to make them sound fluent. Since Kondo’s (2000) finding is not yet conclusive, whether [j] in [CjʉC̥] blocks HVD or not by linguistically naïve native SJ speakers is still unknown. To fill in this gap, I conducted experiment 4.

## 3.2 Experiment 4: HVD in [CjʉC] contexts

### *3.2.1 Participants, stimuli and procedure*

The participants were the same group that participated in Experiment 1. The target morae were the [Cjʉ] forming [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ], the [Cʉ] counterparts forming [kʉ, pʉ, tʉ, ɸʉ] (since *[hʉ] is phonotactically prohibited, [hjʉ] lacks its counterpart). Note that [tjʉ, ɸjʉ] and [tʉ] occur only in recent loanwords, while [kʉ, pʉ] occur in native words, so a word stratum difference may affect the results (see Table 8 below). I also included other [Cʉ] morae written with two kana letters, specifically [ɕʉ, tɕʉ] (シュ, チュ), in order to examine a possibility that blocking of HVD is not due to pure phonology, but due to orthographic influence. I also compared both [Ci] morae that occur in native Japanese words, specifically [hi, tɕi], and [Ci] morae that occur only in recent loanwords, specifically [ti, ɸi], in order to examine a word stratum (or orthography) influence. These target morae are embedded in four-morae nonsense compound words. Again, the first element is [_tá] when the target consonant is non-coronal and [_ká] when the target consonant is coronal to avoid homorganicity. The second element is of a real morpheme, [boɕi] 'star', [baɕi] 'bridge', or [batɕi][8] 'bee' starting with a voiced consonant to avoid a slight possibility of devoicing in the preceding vowel. The [Cjʉ] targets and their [Cʉ] counterparts are followed by the same second element to be consistent in the environment: for example, [kjʉtá-boɕi] versus [kʉtá-boɕi], and [ɸjʉtá-batɕi] versus [ɸʉtá-batɕi]. There were also six other four-mora nonsense words as distracters. The participants were instructed to put the phonemic pitch accent nucleus on the second mora. Since the default place of the accent nucleus in this type of compound words is the last syllable of the first element, the pitch accent pattern of these stimuli should be natural to native SJ speakers. These nonsense words were embedded in the same sentence as the one used in Experiment 1, *watashitachiwa _ desu* 'We are _'. *Watashitachi,* where the underlined /i/ is highly likely devoiced/deleted, could make the participants feel natural to devoice the target morae. The participants were asked to read each sentence in natural speed three times.

This experiment and Experiment 1 in §2.3.1 were done together, and the sentences in these two experiments were shuffled in random order on the same sentence list, so that the sentences in one experiment were distracters for the other experiment (see Appendix for the actual stimuli).

[8] [boɕi], [baɕi], and [batɕi] are underlyingly /hoɕi/, /haɕi/, and /hatɕi/ respectively. In compound words, the first consonants become voiced due to the *rendaku*, or sequential voicing phenomenon.

*3.2.2 Data analysis*

The number of occurrence of HVD in the target was counted. All the three repetitions in each stimulus sentence were analyzed, so that each target mora was pronounced 102 times (3 repetitions × 34 participants).

*3.2.3 Results and discussion*

Overall, 7 out of 34 participants devoiced all the target vowels in all the three repetitions, and the others did not devoice them at least once. Table 8 shows the results.

| Target [C(j)ʉ] | [kʉ] | [kjʉ] | [pʉ] | [pjʉ] | [ɸʉ] | [ɸjʉ] | [tʉ] | [tjʉ] | [hjʉ] |
|---|---|---|---|---|---|---|---|---|---|
| Word stratum[a] | N/S | N/S | N/S | N/S | N/S | Fo | Fo | Fo | N/S |
| % of HVD[b] | 91 | 75 | 84 | 65 | 98 | 66 | 65 | 63 | 67 |
| p-value[c] | *p=0.011 | | *p=0.002 | | *p<0.001 | | p=0.571 | | |

| Target [Ci] | [ti] | [ɸi] | [hi] | [ɕi] |
|---|---|---|---|---|
| Word stratum | Fo | Fo | N/S | N/S |
| % of HVD | 73 | 75 | 94 | 95 |

| Target [Cʉ] | [ɕʉ] | [tɕʉ] |
|---|---|---|
| Word stratum | N/S | N/S |
| % of HVD | 61 | 87 |

*Table 8. The number of high vowel devoicing*

*Note* .[a] N/S (native/Sino)=occurring in native Japanese and Sino-Japanese words and in recent loanwords as well; Fo (foreign)=occurring in recent foreign loanwords but not in Sino-Japanese and native vocabulary. Orange shaded boxes are morae only in recent foreign loanwords. [b] % of HVD is the percentage of vowel devoicing out of 102 tokens. [c] p-values were calculated by the two-tailed paired t-test. Asterisk* indicates a significant difference between [Cʉ] and [Cjʉ] at the $p<0.05$ level.

Between [Cjʉ]s and their [Cʉ] counterparts, [Cjʉ]s were consistently less frequently devoiced, although [tjʉ] and [tʉ], both of which are loanword morae, were not significantly different. This suggests that [j] blocks HVD to some extent. However, if I look at other morae, the loanword CV morae [ti, ɸi] were devoiced less frequently than native/Sino morae [hi, ɕi].

Frequency of HVD may be related to frequency of occurrence in the Japanese vocabulary. According to Takayama (2003), occurrences of [Cj]s are relatively few in native Japanese words compared to Sino-Japanese words, recent loanwords, and onomatopoeias. In addition, Hizume (2003) points out asymmetrical distribution of long and short vowels in Sino-Japanese words,

where [Cj]s generally precede a phonemically contrastive long /ʉː/ rather than short /ʉ/. This means that [kjʉ, pjʉ, hjʉ] (with a short vowel) are infrequent. Also, according to Otake, Hatano, Cutler, and Mehler (1993), over 70% of morae in corpora of Japanese speech are [CV]s (and the rest are [CjV]s, [C]s, [V]s, and the second half of a long vowel). In consideration of these studies, my results could also be interpreted that infrequently occurring morae [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ, tʉ, ti, ɸi] may be less likely to be devoiced regardless of the mora structures.

Another possible interpretation is that *Yōon* [CjV]s and [CV]s, as well as recent loanword [CV]s, are written with two letters of the Japanese *Kana* syllabary, while native [CV]s are written with one letter as shown from (3) to (6):

(3) Native [CV]s [kʉ, ɸʉ, pʉ, hi, ʨi] are ク, フ, プ, ヒ, チ, respectively.

(4) [CjV]s [kjʉ, pjʉ, tjʉ, ɸjʉ, hjʉ] are キュ, ピュ, テュ, フュ, ヒュ.

(5) Loanword [CV]s [tʉ, ti, ɸi] are トゥ, ティ, フィ.

(6) *Yōon* [CV]s [ɕʉ, ʨʉ] are シュ, チュ.

Orthography might have affected the participants' productions. However, there is a case that orthography cannot explain; two-letter [ʨʉ] チュ is slightly more frequently devoiced than one-letter [pʉ] プ (87% vs. 84%). In other words, orthography does not explain the whole story.

As for [ɕʉ] and [ʨʉ], these [CV] morae were arguablly developed from Sino-Japanese words and are written with two *kana* letters (シュ, チュ) respectively. Even though both are [CV] morae with the same phonological status in terms of word stratum, [ɕʉ] was significantly ($p < 0.001$) less frequently devoiced. Interestingly HVD in [ɕʉ] was even less frequent than that in the loanword [Cjʉ] morae [ɸjʉ] and [tjʉ] although the differences were slight. All this indicates that there are many factors involved in HVD, which I will not discuss in this paper.

To sum up, /j/ in /CjʉC/ could be one of the factors of HVD blocking, but since there are other possible factors, further research is still needed.

Meanwhile, although more than 60% of the occurrences of /ʉ/ in /CjʉC/ were devoiced, it does not reject the complex onset hypothesis. For example, English /l/ and /ɹ/ after voiceless consonants are allophonically devoiced as in 'play' [plej] and 'tray' [tɹej] (McMahon, 2002, p. 65). Likewise, allophonic devoicing of /j/ in /CjʉC/ can also be expected in the same logic. This feeding relationship is shown in (7) below.

(7) Underlying representation /ÇjʉÇ/

↓

Voicing assimilation of [j] [Ç̥jʉÇ]

↓

High vowel devoicing [Ç̥j̥ʉ̥Ç]

Another remaining issue is that in my review of literature, there are no studies about whether /i, ʉ/ adjacent to a voiced consonant are in fact never devoiced/deleted or can be devoiced/deleted. Based on my Japanese colleague's and my informal observation, /ʉ/ in words like /sʉ̲mappʉ/ (a Japanese boy band name), in which /ʉ/ is followed by /m/, can be devoiced/deleted at least by some SJ speakers in natural speech. The general assumption that /i, ʉ/ adjacent to a voiced consonant are not devoiced/deleted may need to be reexamined.

### *3.2.4 Conclusion regarding HVD in [CjʉC] contexts*

To conclude, [j] in [CjʉC] could be one of many factors of HVD blocking, and this provides some support to the complex onset /Cj/ hypothesis. However, this may also be due to some other factors, such as infrequent occurrences or orthographic influences. Therefore, the results of my experiment do not strongly support the /Cj/ hypothesis. Meanwhile, the fact that more than 60% of /ʉ/ in /CjʉC/ was devoiced does not reject the /Cj/ hypothesis. Although this experiment is not conclusive, it is still noteworthy that this is the first study of observing HVD patterns in [Cjʉ] morae as well as loanword [Cʉ, Ci] morae.

## 4 Discussion

### 4.1 The number of segments and native speaker intuitions

Japanese has a word whose pronunciation is notoriously difficult even for trained native speakers, きゃりーぱみゅぱみゅ /kjaɾi:pámjʉpamjʉ/ 'Kyary Pamyu Pamyu (a Japanese fashion model and singer).' Although the repetition of bilabial onsets makes one of the reasons for the difficulty in articulation, the major difficulty lies in producing the /CjV.CV.CjV.CV/ structure while maintaining the mora-timed rhythm. Native SJ speakers would be able to pronounce /pámʉpamʉ/ much faster than /pámjʉpamjʉ/. This can be explained by analyzing the /mjʉ/ mora as having three segments, which therefore requires a longer time than two or one-segment morae (i.e. /CV/, /V/, and /C/) in articulation. Likewise, in mimetics, only one [Cj] is permitted in one root, and this is called monopalatality: for example, [pjoko-pjoko] 'flip-flap' but not *[pjokjo-pjokjo] (Mester & Itō, 1989). Again, according to native SJ speaker intuitions, this constraint may be in part because a three-segment /CjV/ mora is phonetically so long that consecutive CjV morae greatly disturb the mora-timed rhythm. Such disturbance is not tolerated in SJ, in which the mora has an important role. Compare the three nonsense words written in Japanese romanization (in which the letter <y> corresponds to the glide [j]) in (8):

(8) (a) namanama (b) nyamanyama (c) nyamyanyamya

According to the opinions from some of the participants of the experiment above, native SJ speakers felt that (8, c) with consecutive [CjV] morae was unpronounceable in natural speed, and that (8, b) was pronounceable in natural speed but not as fast as (8, a). Again, my interpretation is that /CjVCjV/ takes longer than /CVCV/ because the former contains six segments while the latter contains four. In contrast, when I asked a few of the Russian participants to pronounce these three with palatalized/soft consonants, they commented that all the three were easily pronounceable in natural speed. Moreover, reportedly one of the Russian participants had studied Japanese for around three months by the experiment, and he commented that according to his impression, the Japanese [Cj]s were more similar to the Russian /Cj/ clusters than the Russian palatalized /C$^{j}$/s. All this suggests that the syllables in (8, c) behave very differently in Russian and SJ. For Russian speakers, (8, b) and (8, c) are phonologically /n$^{j}$aman$^{j}$ama/ and /n$^{j}$am$^{j}$an$^{j}$am$^{j}$a/ respectively in which /n$^{j}$/ and /m$^{j}$/ are single consonants just like /n/ and /m/.

Likewise, SJ consecutive CV morae with palatal consonants are pronounceable in natural speed, as seen in (9).

(9) (a) /ʦasaʦasa/ (b) /ʨasaʨasa/ (c) /ʨaɕaʨaɕa/

The SJ participants felt that although (9, c) has consecutive palatal (or nonanteriorized) consonants, it was completely pronounceable in natural speed, or that (9, c) might take slightly shorter time than (9, b) due to alternative occurrences of two different places of articulation in (9, b), and shorter time than even (9, a) due to infrequent occurrence of [ʦa] in SJ. The reason that all the three nonsense words in (9) are pronounceable in natural speed is because /ʦa, sa, ʨa, ɕa/ are all two-segment CV morae. This suggests that consecutive SJ /nja/ or /mja/ morae take longer time not because palatal consonants take longer articulation time, but because /nj/ or /mj/ consist of two segments. Such native- speaker intuitions can support the results of the experiments above.

### 4.2 Existence of more plausible Japanese [C$^{j}$]-[C] contrasts

If a language has palatalized consonants, their corresponding plain consonants tend to be velarized for the sake of perceptually clear distinctiveness, as seen in Russian or Irish (Padgett, 2003a). Such contrasts are regarded as [back] contrasts. In this sense, SJ potentially has more plausible [back] contrasts (or [C$^{j}$]-[C$^{ɣ}$] contrasts) than the [CjV] versus [CV] contrasts, i.e. the /CiC/ versus /CɯC/ contrasts (phonetically [C$^{j}$C] vs. [C$^{ɣ}$C] (or [C$^{w}$C]), C = voiceless consonant) involved in the aforementioned /i, ɯ/ devoicing/deletion. Historically, in the contrastive palatalized consonants in Russian, for example, allophonically palatalized consonants in Old Russian or Old East Slavic were later phonologized (Padgett, 2003a). More specifically, one of the lax vowels called 'jer', i.e. [ɪ], caused palatalization of the preceding consonant, as in [dɑnʲɪ] 'tribute'. Even after the famous loss of jers in certain positions, the preceding consonants still stayed palatalized, as in [dɑnʲɪ] → [dɑnʲ] (Padgett, 2003a, p. 307). Likewise, in

Irish, contrastive consonant palatalization was conditioned by the following high and mid front vowels before the loss of vowels in final and interior syllables (Greene, 1973). Compared with the history of contrastive palatalization in these languages, the potential palatalized-velarized (or labialized) contrasts in SJ are the following cases; consonants are palatalized by the following [i] or velarized (or labialized) by the following [ʉ], and even where [i] and [ʉ] are phonetically deleted due to the aforementioned High Vowel Devoicing/Deletion (HVD), palatalization and velarization/labialization of the consonants still remain. Examples are shown in (10) and (11) (examples from Vance, 2008, pp. 209-210).

(10) [kʲ] vs. [kˠ/kʷ]: /kiɕo:/ 'weather' vs. /kʉɕo:/ 'bitter smile'
[kʲɕo:] vs. [kˠɕo:] or [kʷɕo:]

(11) [ɕʲ] vs. [ɕˠ/ɕʷ]: /ɕitai/ 'dead body' vs. /ɕʉtai/ 'subject'
[ɕʲtai] vs. [ɕˠtai] or [ɕʷtai]

In the cases of (10) and (11), /kiC/ versus /kʉC/ and /ɕiC/ versus /ɕʉC/ are phonetically [kʲC] versus [kˠC] (or [kʷC]) and [ɕʲC] versus [ɕˠC] (or [ɕʷC]). These cases in SJ behave much like palatalized-velarized consonants in Russian since these contrasts are caused by high vowel deletion compared to [CjV]s with no velarized counterparts. Moreover, unlike the aforementioned examples like /pámʉpamʉ/ versus /pámjʉpamjʉ/, native SJ speakers would be able to pronounce both [kʲɕo:] and [kˠɕo:] (or [kʷɕo:]) equally easily and quickly, and the same is true for [ɕʲtai] and [ɕˠtai] (or [ɕʷtai]) as well. This is because underlyingly both [Cʲ] and [Cˠ(ʷ)] consist of the same number of segments. In this sense, SJ potentially has two completely different contrasting systems: palatalized- velarized (or labialized) contrasts and complex-simplex (Cj-C) contrasts, although consonant sequences like [kʲC] versus [kˠ(ʷ)C] are currently not phonologized in SJ.

## 5 Limitations and future research

### 5.1 Limitations about the participants

Because both Japanese-speaking and Russian-speaking participants were recruited in Victoria, Canada, many of them were not monolingual. The Russian speakers' length of residence (LOR) in an English speaking country is ranging from 7 months to 4.5 years (median 2.25 years). The Japanese speakers' LOR is ranging from 1 week to 41 years (median 2.25 years). So there may be some L2 influence on their L1s. However, because the stimuli were written with the orthography in their L1s, I believe that they were in their L1 modes as much as possible during the experiments. Moreover, all the participants reported that they had connections with their L1 speaking friends in Canada. The Japanese participant with LOR 41 years, who may present an extreme case, is married to another Japanese participant with an LOR of 34 years. I believe that their regular

contact with their L1 speakers can help reduce the L2 influence on their L1s. Another limitation is that the number of Russian speakers is not large.

### 5.2 Complex onsets /Cj/ or rising diphthongs /iV/

Another possible interpretation of the longer duration of [CjV] is that [j] in Japanese [CjV] belongs to the nucleus rather than the onset, that is, [j] is part of a rising (in sonority) diphthong /iV/. The results of this present study cannot judge whether [j] is part of a complex onset or a complex nucleus. This will be examined in future research.

## 6 Conclusion

I found that Standard Japanese (SJ) [CjV]s, specifically, [gja, nja, mja, bja, ɾja], were highly significantly longer than the [CV] counterparts [ga, na, ma, ba, ɾa], whereas there was no significant difference between the alveo-palatal [dʑa] and the non-palatal [dza]. This indicates that there was no evidence that palatal consonants are longer than their non-palatal counterparts, suggesting that the longer duration of SJ [Cj]s can be better explained by treating [Cj]s as consisting of two segments /CC/s rather than as palatalized consonants /C$^{j}$/s. To further examine the duration of palatalized consonants, I compared Russian uncontroversial /C$^{j}$/s [n$^{j}$a, m$^{j}$a, b$^{j}$a, ɾ$^{j}$a] with the plain/velarized counterparts [na, ma, ba, ɾa]. In unstressed syllables, there was no significant differences between [C$^{j}$V]s and the [CV] counterparts in duration. In stressed syllables, the absolute durations of [m$^{j}$a, b$^{j}$a, ɾ$^{j}$a] were significantly longer than the [ma, ba, ɾa] counterparts, but not as much as the SJ unaccented [mja, bja, ɾja] versus [mja, bja, ɾja]. However, in the ratios between the target [C$^{(j)}$V]s and the following [ka/ta]s, only [b$^{j}$a] was significantly longer than [ba] at the $p<0.05$ level, but notably only by 7%. This is much less noticeable compared to the unaccented SJ [bja], which was significantly longer than [ba] with p=2E-37 by 20%. Therefore, I conclude that in terms of duration, SJ [Cj]s behave like consonant clusters, i.e. /CC/ or /Cj/, as opposed to Russian /C$^{j}$V/s. At least, there is no evidence that SJ [Cj]s are single palatalized consonants. As for High Vowel Devoicing/Deletion (HVD) patterns, I find that [j] in [Cj̥C] can be one of the factors of HVD blocking, but HVD blocking observed in my experiment may also be because infrequent morae are less frequently devoiced, or two-letter morae are less frequently devoiced. The results do not strongly support the /Cj/ hypothesis. Conversely, the fact that /u̥/ in /Cj̥C/ was still devoiced more than 60% does not reject the /Cj/ hypothesis. Since /L/ in /CL/ (L=liquid) onsets in English are also devoiced, devoicing of /j/ in /Cj̥C/ is also expected. Considering the history of Russian and Iris palatalized-velarized contrasts, a potential palatalized-velarized contrasts in Japanese would be /CiC/ versus /Cu̥C/ ([C$^{j}$C] versus [C$^{ɣ}$C] or [C$^{w}$C]) involving HVD, /kiɕo:/ ([k$^{j}$ɕo:]) versus /ku̥ɕo:/ ([k$^{ɣ}$ɕo:] or [k$^{w}$ɕo:]), for example. This contrast in Japanese would be more similar to Russian palatalized-velarized contrasts than the Japanese [Cj]-[C] contrasts.

**Acknowledgements**

I am grateful to Dr. Ewa Czaykowska-Higgins for supervising my research on Japanese [Cj], and her suggestion of comparing Japanese and a Slavic language and integrating Japanese linguists' feedback about vowel devoicing. I am also grateful to Drs. Hua Lin and Sonya Bird for their valuable advice on this research. Dr. Hua Lin also proofread and revised parts of this paper. I also thank my colleague Natallia Litvin for her advice regarding Russian phonology and help in making the stimuli. Last but not least, I am very grateful to all the 42 participants of the experiments.

**References**

Fujimoto, M. & Kiritani, S. (2003). Comparison of Vowel Devoicing for Speakers of Tokyo- and Kinki Dialects. *Journal of the Phonetic Society of Japan, 7*(1), 58-69.

Greene, D. (1973). The growth of palatalization in Irish. *Transactions of the Philological Society* 72(1), 127-136.

Grenon, I. (2005). The status of the sound [z] in Japanese. *Revue Langues et Linguistique, 31*, 64–90.

Hall, T. A. (2000). Typological generalizations concerning secondary palatalization. *Lingua, 110*, 1-25.

Hashimoto, T. (1984). A phonological study of "yō-on" on Japanese. *Journal of the Phonetic Society of Japan 175*(4), 1-2.

Hizume, S. (2003). On 茶 as a Kana Expressing Yo-on in Japanese. *Studies in humanities* 39, 114-103.

Hyman, L. M. (2006). Word-prosodic typology. *Phonology*, *23*(02), 225-257.

Kenstowicz, M. J. (1994). *Phonology in generative grammar*. Oxford: Blackwell.

Kida, A. (1998). Syllable and mora. *Journal of the Phonetic Society of Japan 2*(1), 41-49.

Kochetov, A. (2002). *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization.* New York: Routledge.

Kohno, M. (1998). Mora, syllable, and rhythm – a psycholinguistic study. *Journal of the Phonetic Society of Japan 2*(1), 16-24.

Kondo, M. (2000). Vowel Devoicing and Syllable Structure in Japanese. *Japanese/Korean Linguistics, 9*, 125-138.

Labrune, L. (2012). Questioning the universality of the syllable: evidence from Japanese. *Phonology*, *29*(01), 113-152.

Larson-Hall, J. (2004). Predicting perceptual success with segments: a test of Japanese speakers of Russian. *Second Language Research 20*, 32–76.

Litvin, N. (2014). *An Ultrasound Investigation of Secondary Velarization in Russian* (Master's thesis), University of Victoria, BC.

Matsuzaki, H. (1993). A quantitative study of fluctuation in Japanese loanword-phone notation. *Journal of the Department of Japanese, Tohoku University, 3*. 83-94.

McMahon, A. M. (2002). *An introduction to English phonology* (Vol. 22). Edinburgh: Edinburgh University Press.

Mester, R. A., & Itô, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language*, 258-293.

Nogita, A. (2010). *Examination of the [si] and [ʃi] Confusion by Japanese ESL Learners* (Master's thesis), University of Victoria, BC.

Nogita, A. Yamane, N., & Bird, S. (2013). The Japanese unrounded back vowel /ɯ/ is in fact rounded central/front [ʉ̟ - ʏ]. Ultrafest VI.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*(2), 258-278.

Padgett, J. (2003a). The emergence of contrastive palatalization in Russian. *Optimality Theory and language change 56*, 307-335.

Padgett, J. (2003b). Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory*, *21*(1), 39-87.

Padgett, J. (2008). Glides, vowels, and features. *Lingua 118*, 1937–1955

Parker, S. (2012). Sonority distance vs. sonority dispersion – a typological survey. In Steve Parker, editor. *The Sonority Controversy*. Berlin: De Gruyter, 101-65.

Shinohara, S. (2012). *"Miseru Koe" no Tsukurikata [How to make charming voice]*. Tōkyō, Japan: Kōdansha.

Takayama, M. (2003). Issues in Japanese historical phonology. *Journal of the Phonetic Society of Japan. 7*(1), 35-46

Takayama, M. (2005). Asymmetry of medial approximants j and w in Japanese. *Journal of the Phonetic Society of Japan.9*(1), 30-38

Vance, T. J. (2008). *The sound of Japanese*. New York: Cambridge University Press.

Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America*, *109*(3), 1144-1156.

Yamaoka, T. (2008). On the learning of the English /æ/ sound by Japanese speaking learners of English: Explanation of its levels of difficulty and formulation of its teaching principles. *The Bulletin of Japanese Curriculum Research and Development, 30*(4), 49-58.

**Appendix A**

## 無意味な言葉を含む文を、3回ずつ読んで下さい。

無意味な言葉は、アクセントの位置を必ず2番目に。例えば「たかはし（高橋）」は「か」にアクセント。
ただしアクセントを強調し過ぎず、あくまで自然なスピードで読んで下さい。
例文）私たちはノリマチです。×３回

１．私たちはゴ マ モチです。　　２．私たちはギョ ナ マチです 。

３．私たちはチュカボシです。　　４．私たちはガタバコです。

５．私たちはクタボシです。　　６．私たちはフィタバシです。

７．私たちはヒュタバチです　　８．私たちはラカバコです。

９．私たちはレナモチです。　　１０．私たちはヒタバシです。

１１．私たちはニャカバコです。　１２．私たちはジャカバコです。

１３．私たちはピュタボシです。　１４．私たちはティカボシです。

１５．私たちはチカボシです。　　１６．私たちはリャカバコです。

１７．私たちはテュカバシです。　１８．私たちはマタバコです。

１９．私たちはギャタバコです。　２０．私たちはナカバコです。

２１．私たちはフタバチです。　　２２．私たちはフュタバチです。

２３．私たちはキュタボシです。　２４．私たちはプタボシです。

２５．私たちはビャタバコです。　２６．私たちはモラマチです。

２７．私たちはミャタバコです。　２８．私たちはトゥカバシです。

２９．私たちはシュカボシです。　３０．私たちはバタバコです。

| ３１．私たちはザ カ バコです。 | ３２．私たちはニョ サ モチです。 |
|---|---|
| Мы гатабако | Она хотела гатап |
| Мы накабако | Она хотела натап |
| Мы матабако | Она хотела матап |
| Мы батабако | Она хотела батап |
| Мы ракабако | Она хотела гатап |
| Мы гятабако | Она хотела гятап |
| Мы някабако | Она хотела някап |
| Мы мятабако | Она хотела мятап |
| Мы бятабако | Она хотела бятап |
| Мы рякабако | Она хотела рятап |
| Мы гъятабако | Она хотела гъятап |
| Мы нъякабако | Она хотела нъякап |
| Мы мъятабако | Она хотела мъятап |
| Мы бъятабако | Она хотела бъятап |
| Мы ръякабако | Она хотела ръятап |