WPLC

Working Papers of the Linguistics Circle
of the University of Victoria

Connecting graduate & undergraduate research

Vol. 20

Working Papers of the Linguistics Circle of the University of Victoria

– Vol. 20 –

Connecting Graduate & Undergraduate Research

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://web.uvic.ca/~wplc | wplc@uvic.ca

Table of Contents

(*undergraduate contribution)

CONTRIBUTIONS IN EXPERIMENTAL PHONETICS

CONTRIBUTIONS ON CURRENT TRENDS IN LINGUISTICS AND SOCIETY

Acknowledgments

The 20th volume of WPLC would not have been possible without the ongoing support of the faculty, staff, and students of the UVic Linguistics Department as well as the greater community of individuals whose interest in the study of language maintains our spirits and motivates our efforts. To this Linguistics Circle, we are very grateful. We would also like to thank the UVic Graduate Students' Society for their generosity and the UVic Faculty of Graduate Studies for its support of graduate-level research and scholarship.

Preface to our 20th volume

Making the transition into, and out of graduate studies is not always easy. Students just beginning their programs often face a steep learning curve when it comes to acquiring the skills and research habits that will eventually see them to the end of their programs. For those finishing up, and especially for those who continue their careers in academia, a whole new set of skills needs to be acquired.

In that context, we set about putting together Volume 20 of WPLC as a vehicle to help students at both ends of the spectrum: by creating a venue for grads & undergrads alike to showcase their work, and by further tasking current graduate students with guiding their fellow students, via careful peer-review, toward higher-level research skills and educational possibilities. We asked the teaching faculty in the UVic Linguistics Department to recommend to us work from their undergraduate classes that they felt was exceptional, and we then invited those students to take part in the publishing process along with three of our graduate students. Following much revision, as well as guidance from the sponsoring faculty members, we were able to bring together an eclectic mix of research reflective of the diverse areas of inquiry that the study of language entails.

It is our hope that the readers of the articles collected here will benefit as much from their content as their authors, editors, and reviewers benefitted from the process of creating them.

Editorial Committee,
WPLC 20

Editorial committee

Review committee

Faculty reviewers & undergraduate paper sponsors:

Student reviewers:

Morphological awareness in bilingual Inuktitut–English speakers from the perspective of relational spreading

Barbara Jenni
University of Victoria
bje@uvic.ca

In this paper I put forward the idea of relational spreading to conceptualize morphological awareness in code-mixing bilinguals. This idea was developed in response to noting instances of incorporated English bases into Inuktitut words in original empirical data from Allen, Genesee, Fish & Crago (2009). Certain classes of verbal affixes, e.g. light verbs, require the incorporation of an element into the position of the base, but incorporations are not limited to nominals. The base itself is considered to be essential in Inuktitut, but it appears that the language also accepts base ellipsis. I propose that these conditions support the incorporation of English elements into Inuktitut verbal affixes and that speakers are likely aware of these conditions when code-mixing. I argue that during code-mixing, speakers match lemmas from one language with lexemes from another language. For bilingual speakers, the relations formed between lemmas and lexemes across languages contain morphological information such as whether or not the lemma of an Inuktitut base can be matched with an English lexeme.

1 Introduction

The standard assumption is that the left-most element of the Inuktitut word, the base, is essential (e.g. Swift & Allen, 2002; Sammons, 1993). Incorporations, abundant in Inuktitut word formation, are found in the left-most position. Incorporations can be defined as instances of a (usually bare) noun in close association with, or morphologically attached to a verb (Johns, 2007). The following data example is from Nowak (2009); the incorporated bare noun 'aiviq' is bold, the verbal affix –si is underlined.

(1) bare root

aiviq	-<u>si</u>	-liq	-ramnuk
walrus	-<u>come.across</u>	-begin	-1D.CAUS.INTR

'We two suddenly come across a walrus'

The majority of Inuktitut speakers today are bilingual, learning Inuktitut as their first language from birth and English as their second language, usually upon entering school (Allen, 2007). The data in Allen, Genesee, Fish & Crago's (2009) study show that bilingual speakers can incorporate English elements into Inuktitut utterances. The incorporated English element in (2), found in the left-most position, is in italics and the verbal affix –u is underlined (data from Allen et al., 2009).

(2) bare root

maani	-i	-gii	*goalie*	-u	-nia	-rama?
here	-be	-IMP.2SS		-be	-FUT	-CTG.1SS

'Can you be here so I can be the goalie?'

Observations of bilingual speakers incorporating both Inuktitut and English elements alike into Inuktitut verbal affixes pose interesting questions about speakers' morphological awareness. Morphological awareness can be understood as the knowledge or understanding of how words are created and how to apply this knowledge (e.g. Rice, Libben & Derwing, 2002; McBride-Chang et al., 2008). In this paper I argue that code-mixed incorporations in Inuktitut–English bilingual speech can be analyzed in terms of *relational spreading*, a proposed model of speech production that draws on the theoretical approach of relational morphology. Following Winford (2009), who states that code-mixing is the outcome of bilingual speakers' matching lemmas from one language with lexemes from another language, I propose that in the bilingual lexicon, lemmas and lexemes are stored in relational patterns across languages. I use the word *lemma* to refer to the semantic grammatical properties of words (without their sounds). The formed relations are used by speakers when creating code-mixed utterances.

This paper first presents a short introduction to Inuktitut, followed by a discussion of incorporations in general and the occurrence of English incorporations into Inuktitut words (Sections 2 and 3). In Section 4, I discuss morphological awareness of bilingual Inuktitut–English speakers based on my observations on code-mixed incorporations and base ellipsis. A presentation of selected literature on the bilingual lexicon and bilingual lexical access in section 5 will provide the basis for my proposal of *relational spreading.* In Section 6 I will apply this proposed model to English incorporations into Inuktitut utterances as an expression of morphological awareness in bilingual speakers.

2 Preliminaries

2.1 Inuktitut genetic and geographical information

Inuktitut is one dialect of the Inuit languages found across Canada's North. The Inuit languages are a sub-branch of the Eskimo language branch and part of the Eskimo-Aleut language family. Inuktitut is spoken in the Eastern Canadian Arctic. Of the estimated 35,000 speakers, the majority live in Nunavut and Nunavik. Inuktitut speakers are exposed to English through education, media, and community interactions. English monolinguals form a minority in Nunavut, where Inuktitut is the official language, and Inuktitut speakers switch to English when interacting with English speakers. There are now some indications of a stable bilingualism developing among Inuktitut speakers in Nunavut (Allen, 2007).

2.2 Inuktitut word structure

Inuktitut is a polysynthetic language with a rich morphology of affixation. The language is head-marking with SOV word order. Swift & Allen (2002) state that the minimal structure of verbs and nouns canonically consists of a base, followed by an inflectional ending. In principle, the word base is the lexical root of the word and is either a noun (or pronoun) or a verb. The base sits in the left-most position of the word. A minimum of one and up to six (Sammons, 1993) or eight (Crago & Allen, 1999) suffixes can be added to the base. The basic word structure in Inuktitut can be formulated as follows (Cook & Johns, 2009):

(3) [base (affixes) inflection]

Suffixation in Inuktitut expresses both lexical content and grammatical categories. There are over 400 word-internal morphemes and over 900 verbal and 100 nominal cross-referencing inflections (Crago & Allen, 1999).

2.3 Incorporations

Incorporations occur frequently in Inuktitut and are considered to be obligatory (Johns, 2009). Technically, a (bare) noun that is morphologically attached to or in close association with a verb constitutes an instance of (noun) incorporation. However, in Inuktitut, the incorporated element is not limited to a nominal (Johns, 2007; Nowak, 2009). The following data set from Nowak (2009)

illustrates four possible types of incorporated elements (in bold): (4) a bare root, (5) an inflected nominal complex, (6) a pronoun, and (7) a particle.

(4) bare root
aiviq -si -liq -ramnuk
walrus -come.across -begin -1D.CAUS.INTR
'We two suddenly come across a walrus'

(5) infl. nom. comp.
illu -**tinnut** -aq -tunga
house **-1P.POSS.TERM** -move -1S.INTR
'I arrived (at our) home' (= I went home)

(6) pronoun
asi -qaq -nngit -tuq
other -have.INTR -NEG -3S.INTR
'There was no other'

(7) particle
uattiaru -u -qqau -juq
a.little.while.ago -be -a.little.while.ago -3S.INTR
'It happened a little while ago (evidential)'

The incorporation of a particle in (7) appears to conflict with the previous statement that the base is either a noun (pronoun) or a verb. The explanation for this is that incorporations are not limited to nominals.

2.4 Verbal affix versus verb roots

In Inuktitut, both verb roots and verbal affixes exist. The element glossed as an English verb can be a verb root found in the position of the base (italics in 8) or a verbal affix in the position of the first suffixed element immediately following the base (underlined in 9). The data in (8) and (9) are from Cook & Johns (2009).

(8) verb root
hini -liq -nia -haaq -&unga
sleep -begin -future -about.to -APPOS.1S
'Just before I went to sleep'

(9) verbal affix
tuktu -qaq -nia -&a -'man
caribou -exist -future -expressive -CAUS.3S
'Because there might be caribou'

Verbal affixes appear to govern incorporation. They have been identified by Johns (2007) as a group of nine classes of light verbs. Johns (2009) proposes that these light verbs do not meet the requirement of a lexical root (i.e. base) as the left-most element in the word structure. Therefore, an element is incorporated that subsequently fills the otherwise vacant position of the base. Johns' argumentation is situated within the discussion around semantics in Inuktitut word formation. Nowak (2009) suggests that incorporated elements function as lexical arguments within the context of information management. Important for the analysis at hand is that light verbs depend on the pragmatic properties of the incorporated base in order to fully specify their intended meaning (Cook & Johns, 2009). In (9), the verbal affix (light verb) *–qaq* requires the incorporation of the base *tuktu* 'caribou' in order to convey the meaning 'might be caribou'. Nowak (2009) supports Johns' proposal by saying that "incorporated lexical items are strictly obligatory and are governed by the incorporating verbal affix."

In summary, we find that each incorporation in the data above occurs with a light verb: *-si* 'come.across' in (4), *-aq* 'move' in (5), *-qaq* 'have' in (6), and *-u* 'be' in (7). This confirms the importance of the base in the Inuktitut word. In the case of incorporations, the verbal affix, or light verb, requires the presence of the base in order to specify the meaning of the light verb. This inherent structural aspect of Inuktitut appears to also allow for English elements to fill the position of the base, as I will discuss in the following section.

3 Code-mixed incorporation (English–Inuktitut)

One topic of research on bilingual speech is the occurrence of elements of both languages within the same constituent uttered by a bilingual speaker. Different terminology and theoretical and methodological approaches are used to describe this phenomenon (e.g. Muysken, 2000; Winford, 2009). For the purposes of this paper, I will use the terms *bilingual word formation* and *code-mixing* interchangeably to refer to the phenomena of English elements being inserted or incorporated into Inuktitut utterances.

The majority of Inuktitut–English code-mixing instances observed by Allen et al. (2009) are noun insertions (60.2%), followed by verb insertions (31.5%). Based on Johns' (2007) list of light verbs, I have identified 17 instances of incorporations in the 100 code-mixed data samples provided by Allen et al. (17%). The following data set from Allen et al. represents a sample of incorporations of English bases. The incorporated English items are italicized, verbal affixes are underlined.

(10) bare root (N)

maani	-i	-gii	*goalie*	-u	-nia	-rama?
here	-be	-IMP.2SS		-be	-FUT	-CTG.1SS

'Can you be here so I can be the *goalie*?'

(11) bare root (V)

Share	-ruma	-vit?
	-want	-INTER.2SS

'Do you want to *share*?'

(12) infl. nom. comp.

Ball	-ti	-guma	-jait?
	-CAUS	-want	-PAR.2SS.3SO

'Do you want him to *play ball*?'

(13) adjective

nuja	-ti	*funny*	-u	-lir	-mata
hair	-ABS.2SPL		-be	-PRS	-CTG.3PS

'Your hair is *funny*'

The verbal affixes that incorporate the English elements are all light verbs, as per Johns (2007): *-u* 'be' (10 and 13), *-ruma* 'want' (11), *-guma* 'want' (12).

Muysken (2000) suggests that, in terms of insertions in code-mixed speech of bilinguals, agglutinative suffixes are non-selective and that the language of the base can be switched as long as the base "is equivalent in categorical status to an element from the language of the affix" (p. 76). The data in (4) to (7) and (10) to (13) indicate that Inuktitut verbal affixes don't restrict for what element they incorporate categorically (noun, verb, compound, particle, or adjective) or with regards to language (Inuktitut or English).

In the next section, I outline my assumptions on morphological awareness of bilingual Inuktitut–English speakers based on the above observations on code-mixed incorporations, as well as base ellipsis.

4 Aspects of morphological awareness

Morphological awareness refers to the knowledge or understanding of speakers of how words are created and how to apply this knowledge. Rice, Libben & Derwing (2002) have described morphological awareness in bilingual speakers of a polysynthetic language. Their study demonstrated awareness in bilingual Dene–English speakers to morphological constituents in Dene. Also, Muysken (2000) suggests that bilingual speakers *determine* [italics added] categorical status equivalence when code-mixing insertions with agglutinative suffixes. The results

from Rice et al. (2002) coupled with Muysken's suggestion (2000) indicate that bilingual Inuktitut–English speakers are likely aware of the requirement for light verbs to incorporate a base. When code-mixing, these bilingual speakers apply their knowledge by incorporating English bases into Inuktitut verbal affixes.

Another implication of speakers' morphological awareness of the character of verbal affixes and bases is indicated in a study by Swift and Allen (2002) on verb base ellipsis. Swift and Allen (2002) found that elliptical verb constructions, i.e. omissions of the base component in verbs, are an established phenomenon in Inuktitut conversational discourse. The authors state that their observation is significant, considering that the base is assumed to be an essential element of the Inuktitut word. Two examples are presented below (data from Swift & Allen, 2002); the omitted element is in [ ] in the gloss.

(14) Ø -qquuq -mmat
ZBASE -probably -CTG.3SS
'She probably [is coming]'

(15) ataata -it=lu Ø -niaq -gatta
father -ABS.2SS=and ZBASE -TODAY.FUT -CTG.1PS
'With your father we will [go out] today'

Swift & Allen (2002) state that the omitted element may be a base or a base with one or more postbases (i.e., affixes immediately following the base). The authors also observe that the remaining elliptical structures may not or did not begin with what they call derivational postbases. With the exception of *-qaq-* 'have', what Swift & Allen call derivational postbases closely resembles what Johns (2007) proposes as light verbs obligatorily requiring an incorporated base. My assumption then is that base omission before a postbase (affix) other than a light verb is possible, but base omission before a light verb, i.e. verbal affix, is not possible. Speakers appear to distinguish light verbs from other postbases in that light verbs always require that the left-most position be filled. This supports the above notion that speakers are aware of the requirement for light verbs to incorporate a lexical base.

The code-mixing data presented by Allen et al. (2009) shows that English elements only occurred in the left-most position of a word (incorporated or not), but never in a position where the English element would act as a light verb incorporating an Inuktitut element. Code-mixing speakers appear to treat Inuktitut light verbs as elements that can not be replaced by English verbs.

The implication of the above is that speakers appear to be aware that Inuktitut light verbs require a base, or that the left-most position must be filled. Speakers also appear to be aware that this requirement can be met by incorporating an element that completes the intended meaning of the light verb.

Finally, speakers appear to be aware that the light verb is not restricted to incorporating a (bare) Inuktitut noun, but accepts elements from different categories or languages. We may then ask how the morphological awareness of code-mixing bilingual speakers can be conceptualized in order to explain the above observations. I will next describe language processing and lexical access models that account for the use of both languages in code-mixed incorporations.

5 Bilingual lexicon & bilingual lexical access

One of the most influential ideas on speech production in psycholinguistics is the spreading activation model originally introduced by Collins and Loftus in 1975 (Costa, La Heij & Navarrete, 2006; see also Collins & Loftus, 1975). This model proposes that any representation spreads a proportion of its activation to other representations with which it is linked. Subsequent models based on this idea assume three layers of representation in speech production: conceptual (semantic), lexical, and sub-lexical (phonological) (Costa et al., 2006; see also Dell, 1986; Levelt, 1989; Caramazza, 1997). There is agreement among these theories that, within the conceptual level, activation of one concept spreads to multiple other related concepts and that the activation then spreads to the lexical level (Costa et al., 2006). The spreading of the activation between the lexical and sub-lexical representations is either discrete or cascading. Discrete spreading is restricted in that only the one selected lexical representation activates forward to respective sub-lexical representations. Cascading spreading occurs not only from the selected lexical representation, but also from co-activated lexical representations to their respective sub-lexical representation. All theories assume that the activation is fed forward (Costa et al., 2006), and it has been proposed that the activation may also be fed backward (Costa et al., 2006; see also Dell, 1989).

Costa et al. (2006) acknowledge that it is very likely that both languages are activated in parallel, at some level, in bilinguals when they mix codes. I agree that English incorporations into Inuktitut verbal affixes support a concept of parallel co-activation of both languages, as otherwise a “turn-off/on” switch of languages mid-word would be required. Also, I assume that both languages are always active to some degree and that the activation spread is cascading and not discrete. I also agree with Levelt (1993, in Winford, 2009, p. 295) that, at the conceptual level, the message is still preverbal. De Bot (in Winford, 2009, p. 295) suggests that one of the two languages will be selected to control speech output; I assume that the selection may occur anytime before the preverbal message has spread to the lexical representations of the two languages, but no later than that. The non-selected language continues to be active and engaged in the processing to some extent in parallel (de Bot, in Winford, 2009).

If bilingual speakers have access to both languages, what 'facilitates' their choice of one language over the other? Finkbeiner, Gollan & Caramazza (2006) suggest that theories on word production usually assume a process of competitive lexical access. From this assumption follows the prediction that, for monolinguals, the closer two lexical representations are in meaning, the more difficult it is to select the correct one. This however results in the so-called "hard problem" (quotations in original) in bilingual lexical access. In bilingual speech virtually every concept is associated with synonymous lexical nodes, yet bilinguals are able to make correct lexical selections without any difficulty. Finkbeiner et al. (2006) therefore propose that lexical selection is by threshold, not by competition as otherwise assumed. Their differential activation model suggests that the first node that reaches a threshold at an activation level will be the one that gets selected. Selection of the correct target language lexical node(s) is provided by increasing the rate of activation of the target language lexical nodes *relative* (my emphasis) to that of non-target language lexical nodes. Finkbeiner et al. propose that the bilingual speaker's intention to speak in one language and not the other modulates the rate at which the activation accrues over lexical nodes in the target and the non-target language. The authors conclude that language systems are not turned off/on or switched, but the intentions of the speaker may activate one language more strongly than the other. This suggests that both languages are active and that speakers choose which language will be the selected language for speech output, while the other language remains active. This in return supports my assumption stated above that activation is parallel and cascading, rather than discrete.

Speakers use specific criteria to make their choices in language. According to Meuter (2009), multilingual speakers strive towards optimizing their language performance and increasing efficiency continuously. Meuter (2009) also suggests that environmental cues are taken into consideration when negotiating language selection. Allen et al. (2009) observed that most inserted English nouns in code-mixed utterances have a commonly used Inuktitut equivalent and are not restricted to loan words. The authors also noted that the Inuktitut equivalent is typically more awkward than the monomorphemic English term, while inserted verbs are typically simple English verb roots. This indicates that – in incorporations – bilingual speakers, in order to optimize their performance, intentionally choose an English base to avoid Inuktitut equivalents.

It follows that speakers not only choose one or the other language, but also particular elements of either language when code-mixing. Winford (2009) uses an expanded psycholinguistic model to investigate underlying processes of language contact and bilingual speech. The model is based on Levelt's (1993) spreading activation model and adapted by de Bot (1992, 2001) to account for bilingual speech production. Significantly, the FORMULATOR (or lexical stage) is split into two 'aspects': the lexical selection drawing from lemmas in the lexicon, and

the phonological encoding drawing from lexemes or word forms from the lexicon. I conceptualize matching lemmas and lexemes to represent two aspects of the same 'instance' of a language item, like two sides of the same coin. Winford suggests that "substitution [of an element of one language by an element of another language] may be triggered by the fact that the lexemes in question are associated with the same semantic content, in other words, their lemmas overlap at the level of semantic form" (p. 296). Applied to the case of English base incorporations, an English lexeme is matched with an Inuktitut lemma.

To review, the speech production process suggested to facilitate English base incorporations so far is based on a spreading activation model. Starting with preverbal messages, representations spread in a cascading fashion. Depending on the speaker's intention and choice, one of the two languages activated is selected for speech output, while both languages continuously can be more or less activated relative to the other language. Substitutions may occur by matching the lemma of the selected language with the word form of the language more activated (relative to the other). The final piece to integrate is the speakers' morphological base awareness. More specifically, how can we conceptualize speakers' apparent knowledge about the requirements of a base in first position before an Inuktitut light verb? The proposal discussed in the next section draws on the idea of relational morphology and suggests that, in the bilingual lexicon, lemmas and lexemes of both languages form relations.

6 Relational spreading between lemmas and lexemes

As discussed earlier, the general assumption is that the base of the Inuktitut word is strictly required. Actual utterances from Inuktitut speakers suggest, however, that the base can be omitted, as long as the remaining structure does not begin with what appears to be a light verb. Before light verbs, the incorporation of a base is required and does not appear to be restricted to nominals. In code-mixed speech, Inuktitut appears to allow for incorporations of English bases into Inuktitut verbal affixes. From these observations I argue that, at least with incorporations, the requirement that the left-most position be filled with an element that qualifies as a base does not reside exclusively within the base itself, but is an inherent property of the base *in relation* to the occurrence of a light verb.

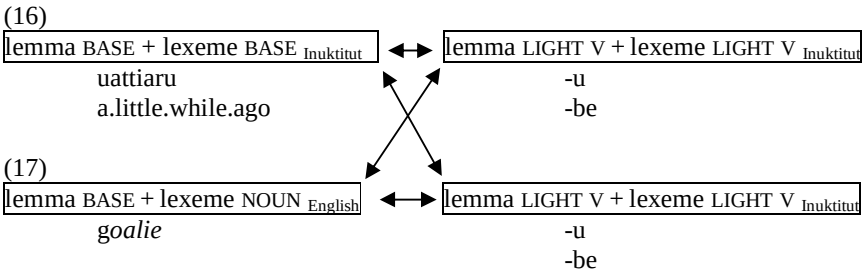
This proposition is in line with the concept of gradient morphology, as discussed by Hay & Baayen (2005). Hay & Baayen have reviewed research supporting the notion that relations between parts/elements are more important than parts/elements by themselves. The specific relation between bases and light verbs results in the required incorporation of a base (not restricted to nominals or Inuktitut elements) into a light verb. The concept of gradient morphology also

agrees largely with Blevins' word-based morphology (2006), which proposes that grammatical patterns, including their inherent relations, reside in actual word forms and that exemplary paradigms and principle part inventories contain word forms. According to Blevins, the paradigms and principles are not minimal in the sense that they are separable from the actual word form. The requirement for a base to fill the left-most position is not separable from either the base or the light verb, but resides in the actual word-form, i.e. the incorporation.

Relational and word-based approaches assume some sort of full-form storage or permanent lexical units. I suggest that the notion of relation can be extended and applied to the concept of lemmas and lexemes via the approach of *relational spreading*. If we assume that spreading activation is cascading (and possibly fed forward and backward), we can suggest that spreading occurs relationally between representations, including between lemmas and lexemes, and across languages. In the monolingual lexicon, lemmas will always match up with the word form of one language, effectively supporting Blevins' word-based morphology. In the bilingual lexicon, where both languages are activated in parallel, we can assume that relations will form between lemmas and lexemes of both languages.

Not all lemmas and lexemes can form relations though, as the possibility of actual relations to occur is rooted in the patterns of a language. Depending on the language selected for speech output and the activated language (relative to the second language), certain relations become available for speakers to intentionally choose from. In the situation where Inuktitut is selected, these relations include morphological knowledge that Inuktitut light verbs require a base to be incorporated and that the choice of base is not limited across categories or languages. Other relations contain the morphological knowledge that only an Inuktitut light verb can be a verbal affix incorporating a base.

The result of matches between lemmas and lexemes is a network of relations across languages. If we isolate the incorporations in the previously discussed data examples in (7) and (10), this can be formalized as follows (referring back to the metaphor that a lemma and lexeme constitute two sides of one coin, the symbol '+' represents where a lemma and a lexeme match; arrows indicate possible relations; the respective language is in subscript):



The allowed matching of the respective lemmas and lexemes is reflected in the relation between the light verb and the base. The lemma LIGHT VERB matched with the lexeme LIGHT VERB $_{\text{Inuktitut}}$ in (16) relates to the lemma BASE matched with the lexeme BASE $_{\text{Inuktitut}}$ just as well as it relates to the lemma BASE matched with the lexeme NOUN $_{\text{English}}$ in (17). For the left-most position to be filled, the light verb is not limited to incorporating a nominal or an Inuktitut element. That is, in (16) the particle '*uattiaru'* is incorporated while in (17) the English word 'goalie' is. On the other hand, the speech of bilingual Inuktitut–English speakers reveals that the Inuktitut light verb is never replaced by an English element. This means that the lemma LIGHT VERB does not appear to match with any form of lexeme VERB $_{\text{English}}$. With no substitutions in the form of *lemma LIGHT VERB matched with a lexeme VERB $_{\text{English}}$ existing, no subsequent relations are formed with a lemma BASE matched with a lexeme BASE $_{\text{Inuktitut}}$ or lexeme BASE $_{\text{English}}$.

7 Conclusion

Observations of incorporated English elements into Inuktitut verbal affixes have prompted this investigation into the morphological awareness of bilingual speakers. From a psycholinguistic perspective and rooted in the theoretical framework of relational morphology, I have proposed that through cascading spreading activation, or *relational spreading*, the lemmas and lexemes of both languages form relations across languages, and that these relations are available to speakers to make choices in language use when code-mixing. Empirical testing in a psycholinguistic setting is needed next to establish the validity of this proposal for further theoretical exploration.

References

Allen, S. (2007). The future of Inuktitut in the face of majority languages: Bilingualism or language shift? *Applied Psycholinguistics* 28, 515–536.

Allen, S., Genesee, F., Fish S. & Crago, M. (2009). Typological constraints on code-mixing in Inuktitut–English bilingual adults. In: M.A. Mahieu & N. Terris (Eds.), *Variations on polysynthesis*. Amsterdam and Philadelphia, PA: John Benjamins Publishing Co., 273–306.

Blevins, J.P. (2006). Word-based morphology. *J. of Linguistics* 42, 531–573.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14(1), 177–208.

Collins, A.M. & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review* 82(6), 407–428.

Cook, C. & Johns, A. (2009). Determining the semantics of Inuktitut postbases. In: M.A. Mahieu & N. Terris (Eds.), *Variations on polysynthesis.* Amsterdam and Philadelphia, PA: John Benjamins Publishing Co., 149–170.

Costa, A., La Heij, W. & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition* 9(2), 137–151.

Crago, M.B. & Allen, S.E.M. (1999). Acquiring Inuktitut. In: O.L. Taylor & L. Leonard (Eds.), *Language acquisition across North America*. San Diego, CA and London: Singular Publishing Group, Inc., 245–279.

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psycholinguistic Review* 93(3), 283–321.

Finkbeiner, M., Gollan, T. & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition* 9(2), 153–166.

Hay, J. & Baayen, R. (2005). Shifting paradigms: Gradient structure in morphology. *TRENDS in Cognitive Sciences* 9(7), 342–348.

Johns, A. (1999). On the lexical semantics of affixal 'want' in Inuktitut. *International Journal of American Linguistics* 65(2), 176–200.

Johns, A. (2007). Restricting noun incorporation: Root movement. *Natural Language & Linguistic Theory* 25, 535–576.

Johns, A. (2009). Additional facts about noun incorporation (in Inuktitut). *Lingua: International Review of General Linguistics* 119, 185–198.

Level, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

McBride-Chang, C., Tardif, T., Cho, J.R., Shu, H., Fletcher, P. & Stokes, S.F. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics* 29, 437–462.

Meuter, R.F.I. (2009). Language selection and performance optimisation in multilinguals. In: L. Isurin, D. Winford & K. de Bot (Eds.), *Multidisciplinary approaches to code switching.* Amsterdam and Philadelphia, PA: John Benjamins Publishing Company, 28–51.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. United Kingdom: University Press, Cambridge.

Nowak, E. (2009). Arguments and information management in Inuktitut. In: M.A. Mahieu & N. Terris (Eds.), *Variations on polysynthesis.* Amsterdam and Philadelphia, PA: John Benjamins Publishing Company, 201–213.

Rice, S., Libben, G. & Derwing, B. (2002). Morphological representation in an endangered, polysynthetic language. *Brain and Language* 81, 473–486.

Sammons, S. (1993). A structural overview of Inuktitut. *Meta: Translators' Journal* 38(1), 11–17.

Swift, M. & Allen, S. (2002). Verb base ellipsis in Inuktitut conversational discourse. *International Journal of American Linguistics* 68(2), 33–56.

Winford, D. (2009). On the unity of contact phenomena and their underlying mechanisms: The case of borrowing. L. Isurin, D. Winford & K. de Bot (eds.), *Multidisciplinary approaches to code switching.* Amsterdam and Philadelphia, PA: John Benjamins Publishing Company, 279–305.

Appendix A

Abbreviations

All data are represented in the standard English alphabet, except '&' – voiceless, lateral fricative

1S	'first pers. singular'
1D	'first pers. dual'
1P	'first pers. plural'
1PS	'first pers. plural subject'
1SS	'first pers. singular subject'
2SS	'second pers. singular subject'
2DS	'second pers. dual subject'
2SO	'second pers. singular object'
3S	'third pers. singular'
3SS	'third pers. singular subject'
3PS	'third pers. plural subject'
3SO	'third pers. singualr object'
SG	'singular'
S	'subject'
O	'object'
ABS	'absolutive case'
ALL	'alliative case'
APPOS	'appositional mood'
CAUS	'causalis mood'
CTG	'contingent verbal modality'
ERG	'ergative case'
EXPL	'expletive'
FUT	'future'
IMP	'imperative verbal modalis'
INTER	'interrogative verbal modalis'
INTR	'intransitive; single argument'
MOD	'modalis case'
NEG	'negation'
PAR	'participative verbal modalis'
POL	'politeness marker'
POSS	'possessive'
PRS	'present tense'
TERM	'terminalis'
V_t	'transitive verb'
V_i	'intransitive verb'
ZBASE	'zero base'

Fixed segmentism and plural infixation in Nuučaan'uł

Mike Anthony
University of Victoria
mda@uvic.ca

This paper provides an argument for the superiority of prosodic approaches to morphological analysis, particularly Optimality Theory, over and against linear approaches characterized by the uninterruptibility criterion of word-hood and the Bloomfieldian conception of the morpheme as a minimal meaningful unit. Specifically, this paper examines three cases of fixed segmentism and two cases of plural infixation from the Southern Wakashan language of Nuučaan'uł that prove difficult to reconcile with the linear approach so construed. Developing two alternative OT analyses which rank constraints of syllable and foot alignment over the correspondence of input and output segments (characteristic of the uninterruptibility criterion), this paper demonstrates the comparative success of the OT approach in both predicting and explaining the patterns present in these problematic datasets in terms of the prosodic word of Nuučaan'uł.

1 Introduction

This paper examines the implications of Nuučaan'uł infixation and fixed segmentism for the concept of the morpheme, the word, and morphological analysis. Specifically, this paper demonstrates how these processes challenge two presuppositions of the linear approach to morphology – a concatenative approach broadly characterized by a Bloomfieldian conception of the morpheme as a minimal meaningful unit and an adherence to the uninterruptibility criterion of word-hood (Bloomfield, 1935). While Bloomfield classically defines the morpheme as an irreducibly meaningful segment or set of segments that feature prominently in the derivation of a word, cases of fixed segmentism in Nuučaan'uł, specifically in repetitive suffix-triggered reduplication (data sets (1), (2), and (3) in the second section of this paper), challenge this concept by providing instances of empty formal units that behave like morphemes in terms of derivation, yet do not contribute to the meaning of the word. The uninterruptibility criterion of word-hood holds that, in order for a group of segments to constitute a word, extraneous segments cannot be introduced that interrupt this group of segments (Bauer, 2003, p. 63–4). This criterion is too

strict, however, and it is difficult to reconcile uninterruptibility with cases of infixation in Nuučaan'uł (data sets (4) and (5)).

This paper argues that an approach to morphological analysis that operates on a prosodic model of word-hood (McCarthy, JJ. & Prince, A.S., 1986, 1993, 1995a, 1999) is preferable to the linear approach on the grounds that it can account for the patterns that emerge in the Nuučaan'uł data sets. The analysis is couched within Optimality Theory (hereafter abbreviated OT; see Prince and Smolensky, 1993). OT conceives of grammar in terms of ranked and violable constraints on well-formedness. GEN, an operation within a grammar, generates a number of candidate forms which are then evaluated by the EVAL operation and ruled out to the extent that a candidate violates a higher ranked constraint of markedness (the degree to which a structure is unacceptable) or faithfulness (the degree to which the segments of the output structure correspond to the input)(see McCarthy, J.J. (1994), McCarthy, J.J. & Prince, A.S. (1995a), McCarthy, J.J. & Prince, A.S. (1995b) for further discussion of Optimality Theory). To this end, this paper will provide two novel OT analyses of plural infixation that draw from a formal analysis of the templates for the reduplicative segment and plural infix patterns, and a proposed hierarchy of well-formedness constraints for Nuučaan'uł to argue for the superiority of a prosodic approach to morphology over the linear (Stonham, 2004, and Kim, 2003).

Before engaging in this project, it is worthwhile to lay out genetic and geographical information about Nuučaan'uł, as well as provide an overview of the morphological processes found in the language with specific focus on the processes of fixed segmentism and infixation examined in this paper.

1.1 Genetic information and morphological processes

Nuučaan'uł is a member of the Southern Wakashan branch of the Wakashan language family, related to Makah and Ditidaht (Kim, 2003, p. 1)[1]. The traditional term for the language is “t'aat'aaqsapa” meaning 'speaking true or straight'; Nuučaan'uł itself means 'all along the mountains and sea' and has at least fifteen dialects spread along the Western coast of Vancouver Island, from Brooke's Peninsula to Barkley Sound (Stonham, 2004, p. 10). It is an endangered polysynthetic language spoken by 150 to 200 speakers (Kim, 2003, p. 1). It is a morphologically complex language featuring suffixation, both partial and double reduplication, infixation, clitics, and incorporation. Interestingly, Nuučaan'uł prohibits compounding (Nakayama, 2001).

[1] The political term 'Nuuchahnulth' includes the Ditidaht people, while the linguistic term excludes the Ditidaht language.

As an agglutinative language, Nuučaan'uł has a considerable number of suffixes, which can express grammatical features or concrete lexical meanings. Nuučaan'uł has over four hundred of these lexical suffixes (Nakayama, 2001, p. 18). Prefixes are formed by reduplication, which occurs in a variety of patterns – CVV, CVc, CVt, CVƛ patterns. Reduplication can be of full syllables, involve vowel-shortening, or be accompanied by lengthening as well (Kim, 2003, p. 182 & p. 195). Although Nuučaan'uł does not prohibit open syllables, Kim provides evidence to suggest that there is a violable constraint against codas in reduplicants (Kim, 2003, p.202). Many cases of reduplication are triggered by somatic body part suffixes or activity suffixes. Other reduplicants can express aspect, plurality, as well as derivational meanings. Double reduplication can occur in certain contexts, namely, when both an aspectual or derivational suffix and an inflectional suffix (e.g. the plural or distributive in Nuučaan'uł) require reduplication (Kim, 2003, p. 6, 176). In all other environments, if two suffixes or inflections require reduplication, it is satisfied by a single reduplicant.

Nuučaan'uł has been characterized as a polysynthetic language that relies on lexical or syntactic incorporation rather than compounding. These lexical incorporations can have idiosyncratic meanings. There are also a number of clitics and enclitics in Nuučaan'uł that can attach to most syntactic units. In the literature, there are a number of interesting questions open for investigations, such as what criteria distinguishes lexical from syntactical incorporation, and whether or not the system of inflection as a whole is a form of clisis.

1.2 Overview of the data

The following data set presents five patterns of infixation in Nuučaan'uł. The first three sets demonstrate the concept of fixed segmentism, providing instances where some default segment always accompanies a process of reduplication in order to avoid the emergence of a marked structure, which violates the preferred syllable structure of the prosodic word (see Alderete, J., et al. (1999) for a discussion of fixed segmentism). The final two sets (4) and (5) are morphologically conditioned infixes marking plurality (see Broselow E. and McCarthy J. (1984) for a discussion of templatic infixation).

2 Data

2.1 Phonologically-conditioned fixed segmentism

The -ƛ- segment in (1) is the most common form of fixed segmentism in Nuučaan'uł, co-occurring with the repetitive and durative aspect triggered

reduplication of open monosyllabic bases. Note that in the data below, this prefixing reduplication also lengthens the base and copy. The -*ƛ̓*-segment seems prespecified to occupy the coda position of the reduplicant, possibly in order to satisfy phonological restrictions on the syllable structure of Nuučaan'uł. Kim suggests that this form emerges from the interaction between a faithfulness, alignment, and complex coda constraints. Later in this paper, I draw from this suggestion to provide an OT analysis of this phenomenon (Kim, 2003, p. 221).

(1) -*ƛ̓*- Fixed Segmentism[2]:

a. tii**ƛ̓**tiiya
DUP- **-ƛ̓-** ti -(y)a·
REP- [ø] rub -DUR
'rubbing'

b. čii**ƛ̓**čiiya
DUP- **-ƛ̓-** čii -(y)a·
REP- [ø] pull -DUR
'pulling'

c. kʷii**ƛ̓**kʷiiya
DUP- **-ƛ̓-** kʷi -(y)a·
REP- [ø] file -DUR
'filing'

d. paa**ƛ̓**paaya
DUP- **-ƛ̓-** ṗa -(y)a·
REP- [ø] give potlatch gift -DUR
'potlatching'

The -*c*- segment in (2) is a relatively rare allomorph of the -*ƛ̓*- infix noted above. Stonham suggests that the conditioning factors that license the choice between the -*ƛ̓*- and the -*c*- allomorphs are unclear, given that both occur in the environment of prefixed reduplication on open monosyllabic roots (Stonham, 2004). Examples (2a) through (c) illustrate this similarity in distribution, though (a) and (b) provide evidence to suggest that the selection of the -*c*- segment is conditioned by dissimilation on the grounds that this segment always and only precedes a lateral affricate (Kim, 2003, p. 215).

(2) -*c*- Infixation:

a. ƛ̓'ii**c**ƛ̓'iiya
DUP- **-c-** ƛ̓'i -(y)a·
REP- [ø] shoot -DUR
'shooting'

b. ƛaa**c**ƛaaya
DUP- **-c-** ƛa -(y)a·
REP- [ø] drive wedge -DUR
'wedge driving'

c, haa**c**huuła
DUP- **-c-** haẇił -(y)a·
REP- [ø] display wealth -DUR
'displaying wealth'

[2] Unless otherwise stated, all data cited from (Stonham, 2004).

2.2 Derivation-triggered fixed segmentism

Certain derivational suffixes trigger reduplication in a manner very similar to the cases examined in (2), and likewise require the pre-specified of a *-c-* segment. Three examples are given below, two of which ((b) and (c)) also involve a separate process of lengthening.

(3) *-c-* infix triggered by derivational suffixes:

a. ʔi**c**ʔinksawiʔak maamaati
DUP- **-c-** ʔink^{w} -sawiƛ' -'aƛ maamaati
SUF- [ø] fire -in eye -TEMP bird
'the birds were blinded by the fires'

b. ʔuu**c**ʔuksuptaakʔaƛ
DUP- **-c-** ʔu -suptaał -'aƛ
SUF- [ø] REF -compete (in) -TEMP
'each tries to be the first to'

c. ʔuu**c**ʔuumahsaqhʔi
DUP- **-c-** ʔu -ɱa· -hsa -(q)h =ʔi·
SUF- [ø] REF -as far -at the brink -MW =DEF
'sit at the very edge of the bluff'

2.3 Consequences of fixed segmentism

Bloomfield (1935) defines a morpheme as a "minimal meaningful unit." The above examples (1-3) feature segments, either *-c-* or *-ƛ-*, which behave like morphemes yet do not contribute to the meaning of the word. How can concatenative approaches to morphology account for these seemingly empty morphemes? In order to account for the behaviour of segments that feature in the derivation of a word yet do not contribute to the words meaning, theorists have proposed both the concept of the morphome and the concept of the formative. The term morphome is meant to stand for a family of morphemes sharing either meaning or the same formal segments (Bauer, 2003, p. 335). Formatives, on the other hand, are solely formal units featuring in derivation that nevertheless do not correspond to a morph (Bauer, 2003 p. 330). Supposing that we grant the validity of these concepts, they do not, in themselves, provide an explanation for why these formatives or morphomes are distributed in the manner they are in the data cited above. Before contrasting the merits of these concepts against OT, it will be worthwhile to broaden the data set to include cases of morphologically-

conditioned infixation that challenge the uninterruptibility criterion of word-hood.

2.4 Morphologically-conditioned infixation

The *-t-* infix in (4) marks the plural, and attaches to the coda of the first syllable of a stem of at least two syllables. The first syllable must contain a long vowel, or else it is lengthened (a-d are cases of lengthening). Further, the *-t-* plural only occurs when the first syllable is open, and the second syllable has a sonorant for an onset (in (4)-(a) this sonorant is /w·/, in (b) /ẇ/, in (c) /ṅ/ and in (d) /ʔ/).

(4) *-t-* plural infixation:

a. naa**t**waaýasʔi
naw·as **-t-** -'as =ʔi
sit idly chatting -PL- -outside =DEF
'those who were sitting outside watching'

b. haa**t**ẇiiqƛhqa
haẇiiqƛ **-t-** -(q)h -qa
hungry -PL- -MW -3.SUB
'they are eating hungrily'

c. t'aa**t**ṅeʔis
t'aṅa **-t-** -ʔis
child -PL- -DIM
'several children'

d. haa**t**ʔum
haʔum **-t-**
food -PL-
'(every kind of) fish'

The *-ẏ-* infix in (5) is in complementary distribution with the *-t-* infix in (4). The *-ẏ-* infix occurs as the onset of the second syllable of a stem composed of at least two syllables. It often occurs between two vowels, and is selected for when the consonant following the first syllable's nucleus is an obstruent. (In 5(a) this obstruent is /p/, in (b) /x/, in (c) /s/ and in (d) /ƛ/).

(5) -*ẏ*- plural infixation:

a. čaẏaapac
čapac -ẏ-
canoe -PL-
‘canoes’

b. caẏaaxuk čapac
caaxuk -ẏ- čapac
swift -PL- canoe
‘swift canoes’

c. maẏaasčim
masčim -ẏ-
commoner -PL-
‘commoners’

d. šiẏaaƛ'aqa
šiiƛ -ẏ- -'aqa
move house -PL- -several doing
‘you're all moving’

2.5 Consequences of the plural infix

The plural infix in these cases presents an instance of a meaningful segment that nevertheless violates the criterion of uninterruptibility for word-hood. Bauer defines uninterruptibility as the condition that “extraneous material cannot be introduced into the middle of a word-form” (Bauer, 2003, p. 63). In each of the above cases, however, the plural infix has interrupted the base form. However, in order to justify the intuition that the form the plural infix interrupts is a genuine word we must adjust our conception of word-hood. In the following discussion, I argue that the conditions for word-hood in Nuučaan'uł and the processes of infixation and fixed segmentism are better understood in prosodic terms, that is, in terms of ranked wellformedness, markedness, alignment and faithfulness constraints.

3 Discussion

3.1 Prosodic templates

The fixed segments in (1–3), in conjunction with the cases of plural infixation in (4) and (5), suggest that a prosodic model of word-hood is more applicable to Nuučaan'uł. That is, a model where word-forms are understood in terms of ideal syllable and foot shapes, marked structures and faithfulness constraints. A quick examination of the data suggests that both the reduplication-triggered segments and the plural infixation align to the right of the initial syllable of the iambic foot. Stonham (2004) suggests that the -*t*- infix must be understood templatically, given that it always co-occurs with lengthening and attaches as a coda to the initial syllable (p. 188). Nonetheless, this account does not sufficiently capture the distribution of the -*ẏ*- infix. For this case, Stonham (2004) proposes a template in which the -*ẏ*- infix inserts to the right of a monosyllabic root but

transforms this into a bi-syllabic foot (p. 193). Stonham (2004) further argues that the *-ẏ-* infix is attached as the onset of this lengthened second syllable (p. 193–4). To demonstrate the virtues of this prosodic approach to word-hood over and above the linear approach, the next subsection of this paper generates two OT analyses that serve to explain the emergence of Stonham's templates.

3.2 Optimality Theory analyses

Drawing from Prince and Smolensky (1993), and McCarthy & Prince (1995a, b) I present two OT analyses of the plural infixes assuming the templates proposed by Stonham (2004). It is the main contention of this paper that OT can correctly predict the form of the plural infixation while the linear approach to morphology, albeit bolstered by adding the notion of formatives and morphomes to its conceptual repertoire, cannot. Insofar as prosodic approaches can better explain the phenomena in question, they are to be preferred.

In the following tableaux, $V_1 \rightarrow V$: is an abbreviation for a family of constraints that ban a short vowel in initial position of plural words, Align-Right-1stSyll is the constraint which holds that any infix must align to the right of the first syllable in the foot, Align-Left-2ndSyll is the alignment constraint which holds that the infix must align to the left of the second syllable in the foot, and IO-DEP (input-output dependence) is a faithfulness constraint that bans insertion: every element of the output must have a correspondent in the input (See McCarthy and Prince, 1995, p. 370). For the purposes of demonstrating the superiority of the prosodic word model, I take the Alignment constraints to be exemplary of the constraints suggested by conceiving of the word in a manner informed by direct acquaintance with the preferred syllable and foot structure of the language. Furthermore, I take IO-DEP to be a constraint that expresses the uninterruptibility criterion of word-hood. As a brief summary, in OT, the GEN operation creates a host of candidate word-forms, listed below as (a) through (d). Once the constraints are ranked, the EVAL operation then evaluates candidates according to the constraint hierarchy; word forms that violate higher ranked constraints are eliminated in favour of those which violate lower ranked constraints. Although other theorists support the rankings I have proposed below, I have placed IO-DEP lower than the Alignment constraint on the hypothesis that the prosodic word model can better account for the actual patterning found in the language. This hypothesis appears confirmed in the tableau below. Should we follow the linear approach, the best candidate would be the unacceptable forms **haʔumt* and **čapacẏ*. Ranking prosodic concerns over interruptibility, however, yields the desired results.

Tableau 1. *-t-* infix

	$V_1 \rightarrow V:$	Align-Right-1stSyll	IO-DEP
a. ☞ haatʔum			**
b. haaʔtum		*!	**
c. hatʔum	*!		*
d. haʔumt	*!	*	

Tableau 2. *-y̆-* infix

	$V_2 \rightarrow V:$	Align-Left-2ndSyll	IO-DEP
a. ☞ čay̆aapac			***
b. čaay̆pac	*!	*	***
c. čay̆apac	*!	*	**
d. čapacy̆	*!	*	

4 Conclusion

In this paper, I have shown that the fixed segmentism and plural infixes of Nuučaan'uł provide important cases in which prosodic approaches to morphology, in this case exemplified by Optimality Theory, are met with a much greater degree of success than the traditional linear approach to the problems presented in this data set. I have presented support for the conclusion that the fixed segments (1), (2) and (3) are best accounted for as units required to complete the well-formed prosodic word structure of Nuučaan'uł, and that this template for the desired word structure can also account for the examples of the plural infix in (4) and (5).

References

Alderete, J., Beckman, J., Benua, L., Gnanadesikan, A., McCarthy, J.J. & Urbanczyk, S. (1999). Reduplication with fixed segmentism. *Linguistic Inquiry* 30, 327–364.

Bauer, L. (2003). *Introducing linguistic morphology.* (2nd ed.). Washington, DC: Georgetown University Press.

Bloomfield, L. (1935). *Language*. London: Allen and Unwin.

Broselow, E. & McCarthy, J. (1984). A theory of internal reduplication. *The Linguistic Review* 3, 25–88.

Kim, E.S. (2003). *Theoretical issues in Nuu-chah-nulth phonology and morphology*. Ph.D. dissertation, University of British Columbia.

Nakayama, T. (2001). *Nuuchahnulth (Nootka) morphosyntax*. Berkeley, California: University of California Press.

McCarthy, J.J. (1994). Morphology, nonconcatenative. In: R.E. Asher (Ed.), *The Encyclopedia of language and linguistics.* Oxford: Pergamon, 2598–2600.

McCarthy, J.J. & Prince, A.S. (1986). Prosodic morphology. Ms., University of Massachusetts, Amherst, & Brandeis University, Waltham, MA.

McCarthy, J.J. & Prince, A.S. (1993). Prosodic morphology I: Constraint interaction and satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University. RuCCS-TR-3.

McCarthy, J.J. & Prince, A.S. (1995a). Prosodic morphology. In: J.A. Goldsmith (Ed.), *The handbook of phonological theory*. Oxford and Cambridge MA: Blackwell, 318–366.

McCarthy, J.J. & Prince, A.S. (1995b). Faithfulness and reduplicative identity. In: J.N. Beckman, L. Walsh Dickey & S. Urbancyzk (Eds.), *University of Massachusetts Occasional Paper 18, Papers in Optimality Theory,* 249–384.

McCarthy, J.J. & Prince, A.S. (1999). Faithfulness and identity in prosodic morphology. In: R. Kager, H. van der Hulst, & W. Zonneveld (Eds.), *The prosody morphology interface*. Cambridge: Cambridge University Press, 218–309.

Smolensky, P. & Prince, A.S. (1993). Optimality Theory: Constraint interaction in generative grammar. *Rutgers University Center for Cognitive Science technical report 2*.

Stonham, J. (2004). *Linguistic theory and complex words: Nuuchahnulth word formation.* New York: Palgrave Macmillan.

Vowel epenthesis in Bengali: An Optimality Theory analysis

Khaled Karim
Ph.D. Candidate, University of Victoria Linguistics
khaledk@uvic.ca

This paper examines the occurrence of epenthetic vowels before and between the initial consonant clusters in Bengali speakers of English, and provides an Optimality Theory (OT) analysis to account for this phenomenon. Native Bengali words disallow initial consonant clusters, and many word-initial consonant clusters in loan words are simplified according to these phonotactics. The maximum syllabic structure is CVC in Bengali and speakers often carry this restriction over to loan words. For example, geram (CV.CVC) instead of gram (CCVC) for the Sanskrit loan word "village", or iskul (VC.CVC) instead of skul (CCVC) for the English word "school" (Kar, 2009). I argue that in rising sonority clusters, a vowel is inserted between the two consonants and in falling sonority clusters (i.e., [s]-stop clusters) the vowel is inserted before the consonant cluster. I also explain that the sonority sequencing constraint SYLLABLE CONTACT treats [s]-stop clusters differently from obstruent-sonorant clusters, and the differing epenthesis pattern can be explained properly if it is considered an effect of SYLLABLE CONTACT – the preference of sonority to fall across a syllable boundary, which was proposed by Murray and Venneman (1983) and also supported by Gouskova (2001). With tableaux, I demonstrate that the epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali and the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). I also demonstrate that the constraint that prefers epenthesis before the [s]-stop cluster is CONTIG-IO (Kager, 1999). Furthermore, I propose that apart from SYLLABLE CONTACT, two other constraints *OO and *OR can also account for the vowel epenthesis in Bengali.

1 Introduction

In this paper, I examine the occurrence of epenthetic vowels before and between the initial consonant clusters in Bengali speakers of English. Native Bengali words disallow initial consonant clusters and many word-initial consonantal clusters in loan words are simplified according to these phonotactics. The maximum syllabic structure is CVC and speakers often carry this restriction over

to loan words. For example, *geram* (CV.CVC) instead of *gram* (CCVC) for the Sanskrit loan word "village" or *iskul* (VC.CVC) instead of *skul* (CCVC) for the English word "school" (Kar, 2009). I argue that this epenthesis is sensitive to sonority. Clusters that rise in sonority are broken up by an epenthetic vowel, and clusters that fall in sonority are resolved by placing an epenthetic vowel before the cluster. In addition, I find that obstruent clusters involving [s] pattern differently from obstruent-sonorant clusters.

The structure of the paper is as follows. First, I present notable features of Bengali language and two sets of data for two vowel epenthesis patterns in Bengali. I then provide an analysis of the data within the framework of Optimality Theory (Prince and Smolensky, 1993). Finally, I demonstrate the constraints in tableaux and present my alternative analysis and conclusion.

2 Notable features of Bengali

Bengali is an eastern Indo-European language with approximately 211 million speakers in Bangladesh and the Indian state of West Bengal. Bengali emerged as a new Indo-Aryan language by 900–1000 AD through Magadhi Apabrangsa and Abahatta, two stages of Magadhi Prakrit (600 BC – 600 AD), along with two other Indo-Aryan languages, Oriya and Assamese (Chatterji, 1926). Until the 14th century, there was little linguistic difference between Bengali and Assamese. Bengali has two literary styles: one is called *Sadhubhasa* (elegant language) and the other *Chaltibhasa* (current language). The former is the traditional literary style based on Middle Bengali of the sixteenth century, while the later is a 20th century creation and is based on the speech of educated people in Calcutta (Sahidullah, 1965). The differences between the two styles are not huge and involve mainly forms of pronouns and verb conjugations (Sahidullah, 1965).

The Bengali alphabet is a syllabic alphabet in which consonants all have an inherent vowel. Vowels can be written as independent letters, or by using a variety of diacritical marks which are written above, below, before or after the consonant they belong to. Word order in Bengali is SOV (Kar, 2009), for example, *ami (I) bhat (rice) khai (eat)* instead of “I eat rice” in English.

3 The problem and related data

The restrictions on word-initial consonant clusters in native Bengali carry over to the pronunciation of English words by Bengali speakers learning English as a foreign language. These learners use a strategy of vowel epenthesis to break up initial consonant clusters.

Sometimes vowel epenthesis occurs between the two consonants of the consonant clusters. For example:

(1)	ENGLISH		BENGALI
a.	Front:	/frʌnt/	/fərʌnt/
b.	Flat:	/flæt/	/fəlæt/
c.	Cream:	/krim/	/kərim/
d.	Group:	/grup/	/gərup/
e.	Floor:	/flɔr/	/fəlor/

(adapted from Islam, 2004)

In some cases epenthesis occurs before the initial consonant clusters. For example:

(2)	ENGLISH		BENGALI
a.	Special:	/'speɪʃl/	/ɪspeɪʃal/
b.	Spain:	/'speɪn/	/ɪspeɪn/
c.	Station:	/'steɪʃn/	/ɪsteʃon/
d.	School:	/sku:l/	/ɪskul/

(adapted from Islam, 2004)

But, when a consonant cluster occurs between two vowels, epenthesis does not occur. For example:

(3)

a.	Astonish:	/əstɒn ɪʃ/
b.	Continue:	/kəntɪn yu/
c.	Monday:	/mʌn deɪ/
d.	April:	/eɪ prəl/

(data source: author)

4 Analysis

The data sets in section 3 illustrate two different strategies to break the consonant clusters. When the words start with obstruent and resonant, the vowel insertion occurs in between obstruent and resonant. And when the words start with obstruent [s] followed by a stop, then epenthesis occurs word initially. In this

section an analysis of the place of epenthetic vowel[1] will be considered for analysis within the framework of Optimality Theory (McCarthy & Prince 1993; Prince & Smolensky, 1993).

In the Optimality Theory (henceforth OT) structure phonological constraints are ranked and violable (Prince & Smolensky, 1993). These constraints are minimally violated by potential surface forms (possible set of candidates) and the one which violates the lowest ranked constraints wins. The seriousness of a violation depends on the hierarchies of constraints and the violations of higher-ranked constraints are most serious. There are two types of constraints: markedness and faithfulness constraints (Prince & Smolensky, 1993). Markedness constraints enforce well-formedness of the output candidate, prohibiting structures that are difficult to produce or comprehend, such as consonant clusters (Prince & Smolensky, 1993). These constraints usually impose restrictions on the occurrence of certain segments. Examples of such markedness constraint are: syllables must not have codas (NOCODA); syllables must have onsets (ONSET); and obstruents must not be voiced (*VDOBS) (Kar, 2009). Faithfulness constraints, on the other hand, impose similarity between input and output. For instance, all morphosyntactic features in the input to be overtly realized in the output (Kar, 2009). Some of the faithfulness constraints are: the output must present all segments present in the input (DEP-IO); elements adjacent in the input must be adjacent in the output (CONTIGUITY); and input segments must have counterparts in the output (MAX-IO) (Kar, 2009).

The data in (1) and in (2) illustrate that there is a restriction against word initial consonant clusters and there is a different epenthesis site for [s]-obstruent clusters. This kind of restriction can be translated into an OT constraint. The constraint is called *$CC_{ONS,}$ which assigns a violation mark to words with consonant clusters (Kager, 1999). For example, outputs like ‘special’ or ‘front’ will not be allowed. An output with word-initial vowel epenthesis to break the consonant cluster is a possible solution. For example: /ɪspeɪʃal/. Another solution is output with vowel epenthesis between consonants of the initial cluster. For example: /fərʌnt/. But these will be possible only at the cost of violation of the faithfulness constraint DEP-IO, which assigns a violation mark to words with epenthetic vowels. Other possible solutions (given example inputs /speɪʃal / and /frʌnt/) are shown overleaf.

[1] The epenthetic vowel, I argue, is a central schwa-like vowel and this schwa can take colour from surrounding segments. But the analysis of the quality or different kinds of vowel is beyond the scope of this paper.

Solutions:	Possible output candidates:
Do nothing (retain 2 adjacent obst. clusters)	[speɪʃal] & [frʌnt]
Appear with one deleted segment	[peɪʃal] & [fʌnt]

Candidates:	Violations:
[speɪʃal] & [frʌnt]	*CC_{ONS}
[peɪʃal] & [fʌnt]	MAX-IO

The constraint MAX-IO assigns a violation mark to words with deleted segments (Kager, 1999). Therefore, outputs like [peɪʃal] and [fʌnt] with one deleted obstruent will be violations of MAX-IO. The constraints, as we see so far, for Bengali data are: *CC_{ONS}, DEP-IO and MAX-IO.

There is an interesting difference in the types of consonants involved in the kinds of consonant clusters which get broken up by internal epenthesis, versus the kinds of consonant clusters which get resolved by initial epenthesis. The first type involves sonorant and obstruent segments and the second type involves obstruents, specifically [s] followed by a stop. The different epenthesis process for '[s]-stop' clusters can be explained by the fact, which Gouskova (2001) correctly observed, that a sonority sequencing constraint such as SYLLABLE CONTACT treats '[s]-stop' clusters differently from obstruent-sonorant clusters. According to Gouskova (2001), the split epenthesis pattern (also evident in Hindi, Sinhalese, Wolof and Uyghur) can be explained properly if it is considered an effect of SYLLABLE CONTACT-the preference of sonority to fall across a syllable boundary, proposed by Murray & Venneman (1983). The epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali but the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). Vowel epenthesis occurs before the cluster whenever the first consonant is of higher sonority than the second consonant (e.g., speɪʃal→ ɪspeɪʃal) and on the other hand, the epenthesis is inside the two initial consonants whenever the first consonant is of lower sonority than the second consonant (e.g., frʌnt→ fərʌnt).

The constraint that prefers epenthesis before the cluster is CONTIGUITY-IO (Kager, 1999). This constraint ensures the epenthesis before the consonants in [s]-obstruent clusters when SYLLABLE CONTACT is not at stake. So, the candidates [sɪpeɪʃal], [fərʌnt] and [əfrʌnt] will have the following violations:

Candidates:	Vioaltions:
[speɪʃal] and [fərʌnt]	CONTIG-IO & DEP-IO
[əfrʌnt]	SYLLABLE CONTACT

While the candidate [əfrʌnt] will have the violation of SYLLABLE CONTACT, the candidate [ɪspeɪʃal] will not violate this constraint.

4.1 Constraint definitions

*CC_{ONS}:	No consonant cluster in the onset.
MAX-IO:	Input segments must have output correspondents (No deletion).
DEP-IO:	No epenthesis.
CONTIGUITY-IO:	No medial epenthesis or deletion of segment.
SYLLABLE CONTACT:	Sonority must not rise across a syllable boundary (Murray & Vennman, 1983; Gouskova, 2001).

4.2 Crucial ranking

The markedness constraint needs to be crucially ranked higher than faithfulness constraint to select an optimal candidate which shows an alternation over other possible candidates which do not. The optimal candidate will violate faithfulness constraints, therefore DEP-IO needs to be ranked lower than *CC_{ONS} to allow epenthesis. This faithfulness constraint DEP-IO is crucially ranked[2] to select the optimal candidate. Another faithfulness constraint CONTIG-IO also needs to be ranked lower than *CC_{ONS} and SYLLABLE CONTACT to determine the optimal candidate. The markedness constraint SYLLABLE CONTACT needs to be ranked above CONTIG-IO to ensure the epenthesis site. Faithfulness constraint MAX-IO needs to be ranked above DEP-IO and CONTIG-IO but below *CC_{ONS} to account for the optimal candidate. But the constraints MAX-IO and SYLLABLE CONTACT[3] are not crucially ranked with respect to each other. So, the ranking of the constraints to account for the Bengali data is as follows in (4):

(4)

*CC_{ONS} >>SYLLABLE CONTACT, MAX-IO>> CONTIG-IO>>DEP-IO.

[2] As it represents the crucial ranking, it needs to be indicated with solid lines in tableaux.
[3] MAX-IO and SYLLABLE CONTACT can be kept in dashed lines as the order of their ranking would provide the same result.

4.3 Tableaux

Tableau 1

/frʌnt/	*CC_{ONS}	SYLLABLE CONTACT	MAX-IO	CONTIG-IO	DEP-IO
a. ☞ fərʌnt				*	*
b. əfrʌnt		*!			*
c. frʌnt	*!				
d. fʌnt			*!		

In Tableau 1, candidate (a) [fərʌnt] is the winning candidate because, although it violates two lower ranked constraints CONTIG-IO and DEP-IO, it satisfies higher ranked constraints. Candidate (b) [əfrʌnt] violates lower ranked candidate DEP-IO but gets ruled out for violating higher ranked constraint SYLLABLE CONTACT. Candidate (c) [frʌnt] loses for violating the higher ranked constraints*CC_{ONS}. The last candidate (d) [fʌnt] also loses for violating MAX-IO, a higher ranked constraint.

Tableau 2

/speɪʃal/	*CC_{ONS}	SYLLABLE CONTACT	MAX-IO	CONTIG-IO	DEP-IO
a. ☞ ɪspeɪʃal					*
b. sɪpeɪʃal				*!	*
c. speɪʃal	*!				
d. peɪʃal			*!		

In Tableau 2, candidate (a) [ɪspeɪʃal] is the winning candidate because it has no fatal violation. Although it violates the lower ranked constraints DEP-IO, it satisfies the higher ranked constraints. Candidate (b) [sɪpeɪʃal] violates CONTIG-IO and DEP-IO but gets ruled out for violating high ranked constraint CONTIG-IO. Candidate (c) [speɪʃal] violates the highest ranking constraint *CC_{ONS} and thus gets ruled out. The last candidate (d) [peɪʃal] violates only MAX-IO but gets ruled out for violating this constraint as this is also a higher ranked constraint.

5 Alternative analysis

As an alternative analysis, I propose that, apart from SYLLABLE CONTACT, two other constraints *OO and *OR[4] can also be used to account for the vowel epenthesis in Bengali. The constraint *OO does not allow two adjacent obstruents in a word, i.e., /speɪʃal/ will not be allowed. The other constraint *OR does not allow obstruents followed by a resonant in a word, therefore /frʌnt/ will not be allowed. It is noteworthy that, CONTIG-IO will also be required to account for the different vowel epenthesis process, as this constraint prefers epenthesis before the consonant clusters (i.e., [s]-obstruent clusters) (Gouskova, 2001). The ranking of the constraint will make sure that the optimal candidates win. The constraints used in the alternative analysis are defined as follows:

*OO:	Two adjacent obstruents are not allowed in a word.
*OR:	Obstruents followed by resonants are not allowed in a word.
CONTIG-IO:	No medial epenthesis or deletion of segment.
*CC_{ONS}:	No consonant cluster in the onset.
MAX-IO:	Input segments must have output correspondents (no deletion).
DEP-IO:	No epenthesis.

5.1 Constraint ranking

The markedness constraint needs to be crucially ranked higher than faithfulness constraint to determine the optimal candidate from the possible candidates. The optimal candidate violates faithfulness constraints, therefore CONTIG-IO and DEP-IO must be ranked lower than *CC. The faithfulness constraint DEP-IO is crucially ranked to determine the optimal candidate. Another faithfulness constraint CONTIG-IO also needs to be ranked lower than *CC_{ONS}, *OR and MAX-IO to determine the optimal candidate. The markedness constraint *OO[5] also needs to be ranked lower to determine the optimal candidate. So, the ranking of the constraints to account for the Bengali data is as follows in (5) overleaf.

[4] These two constraints were proposed by me and used after consultation with Dr. Marion Caldecott, who was the instructor for the course which became the genesis of this paper.

[5] *OO needs to be ranked crucially (solid line) in the case of the candidates in data set 2 to win, as the optimal candidate will violate this constraint. Rest of the constraints can be kept in dashed lines as the order of their ranking would give the same result, i.e., the candidates other than optimal ones would be eliminated if the ranking was different for *OR, CONTIG-IO and MAX-IO; but CONTIG-IO needs to be ranked lower than *CC_{ONS}, *OR and MAX-IO to account for the optimal candidate from dataset 1, as the optimal candidate violates this constraint and it should be kept in solid line.

(5) *$CC_{ONS,}$ *OR, MAX-IO>> CONTIG-IO>>*OO>>DEP-IO

Tableau 3

/frʌnt/	*CC_{ONS}	*OR	MAX-IO	CONTIG-IO	*OO	DEP-IO
a. ☞ fərʌnt				*		*
b. əfrʌnt		*!				*
c. frʌnt	*!	*				
d. fʌnt			*!			

In Tableau 3, candidate (a) [fərʌnt] is the winning candidate because, although it violates two lower ranked constraints CONTIG-IO and DEP-IO, it satisfies higher ranked constraints. Candidate (b) [əfrʌnt] violates the lower ranked DEP-IO and gets ruled out for violating the higher ranked constraint *OR, which is a fatal violation. Candidate (c) [frʌnt] violates two higher ranked constraints*CC_{ONS} and *OR and gets ruled out for violating the highest ranked constraint *CC_{ONS}. The last candidate (d) [fʌnt] has only one violation, i.e., MAX-IO, but gets ruled out as this is a higher ranked constraint.

Tableau 4

/speɪʃal/	*CC_{ONS}	*OR	MAX-IO	CONTIG-IO	*OO	DEP-IO
a.☞ ɪspeɪʃal					*	*
b. sɪpeɪʃal				*!		*
c. speɪʃal	*!				*	
d. peɪʃal			*!			

In Tableau 4, candidate (a) [ɪspeɪʃal] is a winning candidate because, although it violates two lower ranked constraints *OO and *DEP-IO, it satisfies the high ranked constraints. Candidate (b) [sɪpeɪʃal] violates CONTIG-IO and DEP-IO and gets ruled out for violating high ranked constraint CONTIG-IO. Candidate (c) [speɪʃal] violates a lower ranked constraint *OO and also violates the highest ranking constraint *CC and thus gets ruled out. And the last candidate (d) [peɪʃal] violates only MAX-IO but gets ruled out for violating this higher ranked constraint.

6 Conclusion

In this paper, I have provided an OT analysis to account for the vowel epenthesis in Bengali language. I have shown that the primary analysis properly explains the reason why Bengali has a split epenthesis pattern, i.e., the different epenthesis process for [s]-stop clusters than other obstruent clusters (i.e., clusters with obstruent and resonants). I have argued that in rising sonority clusters, a vowel is inserted between the two consonants and in falling sonority clusters (i.e., [s]-stop clusters), the vowel is inserted before the consonant cluster. I have also explained that, the sonority sequencing constraint SYLLABLE CONTACT treats [s]-stop clusters differently from obstruent-sonorant clusters, and the differing epenthesis pattern can be explained properly if it is considered an effect of SYLLABLE CONTACT, the preference of sonority to fall across a syllable boundary, which was proposed by Murray & Venneman (1983) and also supported by Gouskova (2001). With tableaux I have demonstrated that the epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali but the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). I have also demonstrated that the constraint that prefers epenthesis before the [s]-stop cluster is CONTIG-IO (Kager, 1999). Furthermore, I have proposed an alternative analysis, where I have demonstrated with tableaux, that instead of SYLLABLE CONTACT, two other constraints *OO and *OR can also account for vowel epenthesis in Bengali.

Acknowledgements

I am truly grateful to Janet Leonard for providing valuable insights into my data analysis. I am also grateful to my instructor Dr. Marion Caldecott for her valuable suggestions and feedback on several occasions and above all for her care and understanding. Also, thanks to three anonymous reviewers for their useful suggestions.

References

Chatterji, S.K. (1926). *The origin and development of the Bengali language*. Calcutta University, Calcutta.

Gouskova, M. (2004). Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology* 21, 201–250.

Islam, A. (2004). *L1 influence on the spoken English proficiency of Bengali speakers*. C-Essay in English, Högskolan Dalarna University.

Kager, R. (1999). *Optimality Theory.* Cambridge: Cambridge University Press.

Kar, S. (2009). *The syllable structure of Bangla in Optimality Theory and its application to the analysis of verbal inflectional paradigms in distributed morphology.* Ph.D. dissertation, Tübingen University.

McCarthy, J.J. & Prince, A. (1993). Generalised alignment. In: G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993*. Doedrecht: Kluwer, 79–153.

Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint interaction in Generative Grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Shahidullah, M. (1965). *Bangla bhasar itibirtto*. Dhaka: Mowla Brothers.

Learning to read and the development of phonological awareness: Altering our pedagogical approach

Lyra Magloughlin
University of Victoria
lyra@uvic.ca

Despite a century of effort aimed at identifying and implementing effective strategies for teaching reading, children from every demographic, learning in a broad range of environments, continue to demonstrate below grade-level reading performance (e.g. Dechant, 1991; Sherman & Ramsey, 2006; Jacobs, 2008). The research presented in this paper explores the relationship between the development of phonological awareness and the process of learning to read. Several aspects of awareness are discussed, and one developmental model is considered. Two intervention studies demonstrating improved performance following phonological awareness training are examined. Proposed is a move away from labels such as 'developmental dyslexia,' and a shift toward research aimed at providing educators with the tools needed to more adequately meet the developmental needs of struggling readers.

1 Introduction

Unlike language acquisition, the process of learning to read requires more than just exposure to language through social contact. Building the skills necessary for reading demands not only immersion in language and print, but also guided, focused attention and determination on the part of both learner and instructor. As experimental psychologist Steven Pinker (1997) once suggested, "[c]hildren are wired for sound, but print is an optional accessory that must be painstakingly bolted on" (p. ix). In this paper, I will explore the relationship between sound and print. More specifically, I will present evidence to support the notion that children's developmental awareness of language sounds and patterns plays a vital role in their reading success. I will suggest that research aimed at understanding the nature of this role may lead to more effective strategies for teaching reading. I will also suggest that freely assigning labels such as 'dyslexia' warrants more careful consideration. I will argue that children who are struggling with reading

provide us with a tremendous opportunity to rethink conventional pedagogical approaches.

In order to support these contentions, I will first provide background information (§2) in order to position current research in reading. I will then present an overview of the role of phonological awareness in reading (§3) and describe how this awareness involves more than just sounds (§4). Next I will discuss modeling phonological awareness development (§5) and present evidence in support of phonological awareness training and intervention (§6). Finally, I will briefly discuss phonological awareness and dyslexia (§7) and present my conclusion and summary comments (§8).

2 Background

For over a century, researchers have been exploring the mysteries of how our brains learn to read (Quantz, 1897; Huey, 1900; Gates, 1921; Wilson, 1942; Austin & Morrison, 1963; Davis, 1971; Goodman, 1976; Chall, 1983; Ehri, 1995; Jacobs, 2008), with many seeking to identify progressive stages of reading development. In 1995, for example, reading researcher Linnea C. Ehri introduced a model that organized stages of reading into pre-alphabetic, partial alphabetic, full alphabetic and consolidated alphabetic phases. This model provided a flexible framework for researchers attempting to understand the reading process (Beech, 2005). While earlier views have held that instruction in reading should commence when a child demonstrates a certain degree of 'readiness' (see Pearson, 1984, for an historical overview), more recent evidence suggests that the process starts in infancy, and that pre-alphabetic exposure to rhyme and print is a critical part of a child's later reading success (Wagner, Piasta & Torgesen, 2006).

In the mid-1960s a "Great Debate" was waged (Chall, 1967) over how to teach children to read, and in the early 1990s, mainstream media coined the term "Reading Wars" (Vacca, Vacca, et al., 2009) to reflect this ongoing battle. Yet decades of disagreement over reading have not led to the development of a single, wholly effective instructional model. There are researchers who suggest that bottom–up, phoneme-based approaches to teaching reading are critical for helping children make phoneme (sound) to grapheme (symbol) connections (LaBerge & Samuels, 1974; Rayner & Pollatsek, 1989). Others argue for top–down, concept-based approaches, which call on children to identify whole words rather than parts of words (Goodman, 1967; Smith & Goodman, 1971; Goodman, 1996). Still others have advocated for interactive models that engage many different approaches (Perfetti, 1985; Dechant, 1991). These strategies and other variations have been implemented in classrooms across North America, with different instructional models coming in and out of fashion over time. One of the latest of these trends occurred in 2001, when the United States government

enacted legislation calling for the implementation of a phonics-based approach to reading instruction following recommendations made by a government-appointed National Reading Panel (NRP, 2000). Yet, despite a century of effort aimed at identifying and implementing effective teaching strategies, children from every demographic, learning in a broad range of environments, continue to demonstrate below grade-level reading performance (Dechant, 1991; Sherman & Ramsey, 2006; Jacobs, 2008).

3 The role of phonological awareness in reading

Over the past several decades, a large body of evidence has been gathered in support of a reciprocal (Adams, 1990; Bentin & Leshem, 1993; Blachman, 2000), even causal (Bradley & Bryant, 1983; Wagner & Torgeson, 1987; Wagner, Torgeson & Rashotte, 1994; Lundberg, 2009) relationship between children's development of phonological awareness and the process of learning to read. In fact, the term 'phonological awareness' has almost become mainstream and is now frequently referenced in literacy pamphlets, on websites, and in school newsletters as one of the critical components of reading success. But what is phonological awareness and how does it develop? Wagner, Piasta and Torgeson (2006) define phonological awareness as "an awareness of and access to the sound structure of one's oral language" (p. 1114). Pinker (1994) describes phonology as "…the sound patterns of a language, including its inventory of phonemes, [and] how they may be combined to form natural-sounding words…" (p. 480). For the purposes of this paper I will build on Pinker's description in order to define phonological awareness as a conscious understanding of language-specific sound patterns, including how phonemes (sounds) are combined and manipulated to form 'natural-sounding' syllables, rhymes, and words.

In the context of reading, it follows that an awareness of language-specific patterns and combinations may be critical for making important reading connections. Learning to use an alphabetic system, for example, requires early readers to develop an awareness of the connections between phonemes and graphemes. Languages such as German have predictable patterns with one-to-one (phoneme-to-grapheme) correspondences, but languages such as English have complex, often unpredictable, many-to-one (phoneme-to-grapheme and grapheme-to-phoneme) relationships. For English readers, developing an awareness of these unpredictable relationships can be difficult when common English words such as 'do', 'too', 'blue', and 'few' all end in the same phoneme, but are represented by different graphemes. Individuals who experience difficulty with reading frequently have trouble discriminating between phonemes found in the everyday words of their language and tend to perform poorly on blending and

segmenting tasks that require awareness of these language-specific patterns (Siegel & Faux, 1989; Mann, 1993; Oudeans, 2003).

4 More than just sounds

Awareness of the phonological structure of one's language, however, is not limited to audible sounds. It is, perhaps, not surprising that the average deaf individual is only reading at a third- or fourth-grade level by high school graduation (Moores, 1996; Paul, 1998; Gallaudet Research Institute, 1996), given the physical limitations of making an auditory connection between symbols and sounds (Gravenstede, 2009; Paul, 1998). What may be surprising is that profoundly deaf readers who do succeed in reading at or above grade-level seem to exhibit an awareness of phonological patterning during reading. Research in this area (Hanson et al., 1984; Hanson, 1992; Nielsen & Luetke-Stahlman, 2002; Diagle & Armand, 2008; Wang et al., 2008; Aparicio et al., 2009) suggests that although cognitive mapping of symbols to auditory sounds may not be possible for deaf readers, other forms of mapping may be conceivable through exposure to oral speaking, lip reading or 'speechreading', fingerspelling, articulatory feedback, and attentiveness to the ways in which speech sounds are physically articulated in the vocal tract.

Additional evidence in support of the notion that phonological awareness is not limited to audible or acoustic information comes from studies with infants as young as four months old, who are able to discriminate between languages of different rhythmical classes (e.g. stress-timed languages such as English and syllable-timed languages such as French) simply by watching speakers' silent facial movements (Werker & Byers-Heinlein, 2008). As well, some educators now believe that the act of physically articulating sounds while attempting to construct or encode words may have longer-lasting neuronal stability with early readers (Herron, 2008).

Findings such as these require a reshaping of our understanding of phonological awareness to encompass more than just acoustic or auditory information. As the research with profoundly deaf readers suggests, attentiveness to the particular ways in which sounds are physically articulated or visually represented may play an important role during the phoneme–grapheme mapping process – a necessary part of reading proficiency.

5 Modeling phonological awareness development

Interest in understanding and defining stages of phonological awareness is not new, and many developmental models have been proposed over time (e.g. Stanovich et al., 1984; Yopp, 1988; Adams, 1990; Stahl & Murray, 1994; Smith et al., 2007). Although the sequencing of stages tends to vary between models, most identify rhyming, blending, segmenting, deleting, and substitution stages of phonological or sound pattern awareness. In 2007, a group of researchers (Smith, Cassady, Bottomley & Popplewell, 2007) introduced the Standardized Assessment of Phonological Awareness (SAPA) model, which was designed to address some of the gaps and overlaps of earlier models. In order to test SAPA, Cassady, Smith & Putman (2008) developed fourteen discrete tasks that incorporated rhyming, oddity identification, blending, segmenting, phoneme deletion, and substitution. Participants were asked, for example, to blend body-coda or onset-rime segments, and select words in a series with different beginning, middle, or ending sounds.[1] The SAPA tasks were administered longitudinally to participating kindergarten children during the fall, winter and spring of one school year. These discrete tasks enabled Cassady et al. (2008) to measure the sequence of particular aspects of phonological awareness at very specific stages of development. Results strongly supported the researchers' contention that acquisition of phonological awareness occurs in discrete, measurable, developmental stages.

6 Phonological awareness training and intervention

In an attempt to answer the question of whether reading skills could be improved by stimulating phonological awareness, Danish researchers Lundberg, Frost, & Petersen (1988) provided 235 preschool children with 15–20 minutes of phonological awareness training per day, over a period of eight months. The training involved metalinguistic games and exercises, which were designed to help the children develop an awareness of the phonological structure of their language. In Denmark, children do not begin formal reading instruction using an alphabetic script until the age of seven. Following the eight months of training and during their first two years of school, Lundberg, Frost & Petersen tracked the children's reading and spelling progress. Children who had been exposed to explicit phonological training during preschool demonstrated significantly stronger reading and spelling skills during Grades 1 and 2 than did children in the control group, who had received no early phonological awareness training.

[1] While a description of each of these tasks is outside the scope of this paper, a comprehensive list has been included as Appendix A.

Results suggested that phonological awareness training administered prior to formal reading instruction may, in fact, facilitate the process of learning to read.

In another study, Richards & Berninger (2008) scanned the brains of 18 dyslexic and 21 non-dyslexic children at two different times during their performance of an fMRI phoneme mapping task. The first brain scan occurred prior to any intervention, and the second scan occurred after dyslexic participants had received three weeks of phoneme awareness and alphabetic training. The first brain scan for both dyslexic and non-dyslexic children showed very different patterns of fMRI connectivity: the children with dyslexia exhibited greater functional connectivity in bilateral regions of the inferior frontal gyrus, whereas children without dyslexia showed no significant activity in these regions. However, following the second brain scan, which was performed after dyslexic children had received three weeks of phonological training, fMRI connectivity patterns in dyslexic children more closely resembled those of the non-dyslexic children. While the researchers in this study acknowledged that findings are preliminary, they suggested that dyslexic brain connectivity results may be linked to impairment in working memory, and argued that instructional intervention may help children to "overcome" specific temporal deficits (Richards & Berninger, 2008).

Findings such as these are intriguing: if explicit instruction in phonological awareness can lead to improved performance in reading, then a deeper understanding of phonological awareness development could lead to better instructional design.

7 Phonological awareness and dyslexia

Deficits in phonological processing are central to reading difficulties (Ziegler & Goswami, 2005). Since 1887, the term 'dyslexia' has been used to describe individuals who have difficulty reading (Pollak, 2005). The *Oxford English Dictionary* defines dyslexia as "a difficulty in reading due to affection of the brain…word-blindness." Within the reading research community, the term dyslexia is commonly employed and generally classified as either 'acquired' or 'developmental'. Acquired dyslexia is used to refer to individuals who experience difficulty in reading as a result of brain injury or illness, whereas developmental dyslexia is often used categorically to describe individuals who show unexpectedly poor performance in reading. Of critical importance in the classification of developmental dyslexia is that an individual's poor performance in reading is *unexpected*. There is an assumption that individuals with developmental dyslexia do not perform poorly as a result of overt physical or mental impairments, low socioeconomic status, or lack of access to good instruction, but rather due to neurobiological factors that interfere with their

acquisition of sufficient reading skills (Wagner, Piasta & Torgesen, 2006). Developmental dyslexia has been described as a 'disorder', a 'syndrome', a 'disability', and a 'deficit', but regardless of the terms used to describe it, dyslexia generally is considered to be a problem inherent within the individual, rather than a failing on the part of educators. As described above, a number of studies have demonstrated the positive effects of phonological awareness training in addressing certain phonological deficits. The ability to isolate very specific areas of phonological processing difficulties could prove beneficial for those individuals 'diagnosed' with dyslexia.

8 Conclusion

In this paper, I have provided evidence to support the idea that phonological awareness development is central to reading. This idea is generally accepted within the reading research community, yet the debate continues over the efficacy of various instructional models for teaching reading. I propose a move away from this long-standing debate and toward a deeper understanding of the development of phonological awareness.

As many of the studies discussed in this paper suggest, phonological awareness is not limited to the perception of acoustic information, but also involves an awareness of more subtle cues that are produced during the articulation of speech sounds. If profoundly deaf readers and infants are capable of discerning these phonological cues, then hearing children and adults may also be able to access them. Explicit attentiveness to the specific articulatory movements that distinguish one phoneme from another may prove useful in building phonological awareness.

Phonological awareness development models such as SAPA can provide researchers with exciting new avenues for exploring the relationship between awareness and reading. Although existing models do not include aspects of physical articulation as relevant cues for accessing phonological awareness, incorporating this aspect of development into future working models may be an important next step. The evidence pointing to improved reading ability following phonological intervention is encouraging, and tools that enable educators to pinpoint specific areas of developmental difficulty in children who are struggling to read may facilitate the creation of improved instructional materials that succeed in meeting individual learning needs.

Children who struggle with reading face a broad range of challenges that can often be compounded when labels like 'dyslexia,' or 'learning disabled' are used to describe them. As the famous anthropologist Edward Sapir (1929) argued, "[w]e see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of

interpretation" (p. 210). Children who are 'diagnosed' as dyslexic carry with them labels that can influence performance and have long-lasting socio-emotional consequences, as these labels are often introduced at a time when children are working to define self-concept (Pollak, 2005). If we assign labels like dyslexia, we run the risk of approaching each child from the perspective of there being a problem inherent within the child, rather than a failing on our part to adequately meet the child's developmental needs. I propose that struggling readers may provide us with a tremendous opportunity to alter our pedagogical approach. If we can maintain a research focus that is rooted in identifying discrete stages of phonological awareness development, we may be able to more accurately target areas for instructional intervention and more adequately meet the learning needs of children experiencing reading difficulties. This approach seems a promising step in the right direction.

Acknowledgment

I am very grateful for the guidance and support of Dr. Dave McKercher – an exceptional educator and mentor.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Aparicio, M., Demont, E., Gounot, D. & Metz-Lutz, M. (2009). Is there an alternative cerebral network associated with enhanced phonoligical processing in deaf speech-users? An exceptional case. *Scandinavian Journal of Psychology* 50, 445–455.

Austin, M.C. & Morrison, C. (1963). *The first R: The Harvard report on reading in elementary schools*. New York: Macmillan.

Beech, J.R. (2005). Ehri's model of phases of learning to read: A brief critique. *Journal of Research in Reading* 28(1), 50–58.

Bentin, S. & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia* 43(1), 124–148.

Blachman, B.A. (2000). Phonological awareness. In: M.L. Kamil, P.B. Mosenthal, D. Pearson & R. Barr (Eds.), *Handbook of reading research*. Mahwah: NJ: Lawrence Erlbaum Associates.

Bradley, L. & Bryant, P.E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature* 301, 419–421.

Cassady, J.C., Smith, L.L. & Putman, S.M. (2008). Phonological awareness development as a discrete process: Evidence for an integrative model. *Reading Psychology* 29, 508–533.

Chall, J.S. (1967). *Learning to read: The great debate; an inquiry into the science, art, and ideology of old and new methods of teaching children to read*. New York: McGraw Hill.

Chall, J.S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Daigle, D. & Armand, F. (2008). Phonological sensitivity in severely and profoundly deaf readers of French. *Reading and Writing* 21(7), 699–717.

Davis, F.B. (1971). Psychometric research in reading comprehension. In: F. Davis (Ed.), *Literature of research in reading with emphasis on models*. Brunswick, NJ: Rutgers University Press.

Dechant, E. (1991). *Understanding and teaching reading: An interactive model*. Hillsdale, NJ: Lawrence Erlbaum.

Ehri, L.C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading* 18(2), 116–125.

Gallaudet Research Institute. (1996). Stanford achievement test (9th ed.). *Form S: Norms booklet for deaf and hard-of-hearing students. (Including conversions of raw score to scaled score & grade equivalent and age-based percentile ranks for deaf and hard-of-hearing students)*. Washington, DC: Gallaudet University.

Gates, A.I. (1921). An experimental and statistical study of reading and reading

tests. *Journal of Educational Psychology* 12 (September).
Goodman, K.S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist* 6, 126–135.
Goodman, K.S. (1996). *Ken Goodman on reading*. Portsmouth, NH: Heinemann.
Gravenstede, L. (2009). Phonological awareness and decoding skills in deaf adolescents. *Deafness and Education International* 11(4), 171–190.
Hanson, V.L. (1992). Phonology and reading: Evidence from profoundly deaf readers. In: D. Shankweiler & I.Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle*. Michigan: University of Michigan Press, 69–89.
Hanson, V.L., Lieberman I.Y. & Shankweiler, D.J. (1984). Linguistic coding of deaf children in relation to beginning reading success. *Journal of Experimental Child Psychology* 37, 378–393.
Herron, J. (2008). Why phonics teaching must change. *Educational Leadership*, 66(1), 77–81.
Huey, E.B. (1900). On the psychology and physiology of reading. *American Journal of Psychology* 11, 283–302.
Jacobs, V.A. (2008). Adolescent literacy: Putting the crisis in context. *Harvard Educational Review* 78(1), 7–39.
LaBerge, D. & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology* 6, 293–323.
Lundberg, I., Frost, J. & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly* 23(3), 263–284.
Lundberg, I. (2009). Early precursors and enabling skills of reading acquisition. *Scandinavian Journal of Psychology* 50(6), 611–616.
Mann, V.A. (1993). Phoneme awareness and future reading ability. *Journal of Learning Disabilities* 26(4), 259–269.
Moores, D. (1996). *Educating the deaf: Psychology, principles, and practices* (4th ed.). Boston: Houghton-Mifflin.
National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction – Summary report*. Rockville, MD: National Institute of Child Health and Human Development.
Nielsen, D.C. & Luetke-Stahlman, B. (2002). Phonological awareness: One key to the reading proficiency of deaf children. *American Annals of the Deaf* 147(3), 11–19.
Oudeans, M.K. (2003). Integration of letter–sound correspondences and phonological awareness skills of blending and segmenting: A pilot study examining the effects of instructional sequence on word reading for kindergarten children with low phonological awareness. *Learning Disability Quarterly* 26(4), 258–280.

Paul, F. (1998). *Literacy and deafness: The development of reading, writing, and literate thought*. Boston: Allyn & Bacon.

Pearson, P.D. (1984). *Handbook of reading research*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C.A. (1985). *Reading ability.* New York: Oxford University Press.

Pinker, S. (1994). *The language instinct*. New York: Harper-Collins Publishers.

Pinker, S. (1997). Foreword. In: D. McGuinness, *Why our children can't read – And what we can do about it: A scientific revolution in reading*. New York: Simon & Schuster.

Pollak, D. (2005). *Dyslexia, the self and higher education: Learning life histories of students identified as dyslexic.* Sterling, VA: Trentham Books Limited.

Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading.* New York: Prentice-Hall.

Quantz, J.O. (1897). Problems in the psychology of reading. *Psychological Monographs* 2(5).

Richards, T.L. & Berninger, V.W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics* 21, 294–304.

Sapir, E. (1929). The status of linguistics as a science. *Language* 5(4), 207–214.

Sherman, L. & Ramsey, B. (2006). *The reading glitch: How the culture wars have hijacked reading instruction – And what we can do about it*. Lanham, MD: Rowman & Littlefield Education.

Siegel, L.S. & Faux, D. (1989). Acquisition of certain grapheme–phoneme correspondences in normally achieving and disabled readers. *Reading and Writing: An Interdisciplinary Journal* 1, 37–52.

Smith, F. & Goodman, K.S. (1971). On the psycholinguistic method of teaching reading. *The Elementary School Journal* 71(4), 177–181.

Smith, L.L., Cassady, J.C., Bottomley, D. & Popplewell, S. (2007). *The Standardized assessment of phonological awareness*. Muncie, IN: Ball State University.

Stahl, S.A. & Murray, B.A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology* 86, 221–234.

Stanovich, K.E., Cunningham, A.E. & Cramer, B.B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology* 38, 175–190.

Vacca, J.L., Vacca, R.T., Gove, M.K., Burkey, L.C., Lenhart, L.A. & McKeon, C.A. (2009). *Reading and Learning to Read* (7th ed.). Toronto, ON: Pearson Publishing.

Wang, Y., Trezek, B.J, Luckner, J.L. & Paul, P.V. (2008). The role of phonology and phonologically related skills in reading instruction for students who are deaf or hard of hearing. *American Annals of the Deaf* 153(4), 396–407.

Wagner, R.K., Piasta, S.B. & Torgesen, J.K. (2006). Learning to read. In: M.J. Traxler & M.A. Gernsbacher (Eds.), *Handbook of psycholinguistics*. London, UK: Elsevier Inc., 1111–1142.

Wagner, R.K. & Torgeson, J.K. (1987). The nature of phonological awareness and its causal role in the acquisition of reading skills. *Psychological Bulletin* 101, 192–212.

Wagner, R.K., Torgeson, J.K. & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology* 30, 73–87.

Werker, J.F. & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Science* 12(4), 144–152.

Wilson, F.T. (1942). Early Achievement in reading. *The Elementary School Journal* 42(8), 609–615.

Yopp, H.K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterl*y 23, 159–177.

Ziegler, J.C. & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin* 131(1), 3–29.

Appendix

SAPA representative items and abridged instructions (from Cassady et al., 2008, p. 521)

Subtest	Basic Task Requirement	Sample Item(s)
Rhyme recognition	Rhymes are words that sound the same at the end.	Tell me if these words rhyme ape-knee; dip-hip
Rhyme application	Tell me a word that rhymes with:	Star (accept any word or nonsense word that rhymes)
Oddity tasks: Beginning	Listen to the names of these pictures. Tell me which one has a different beginning sound.	duck, door, foot
Oddity tasks: End	Listen to the names of these pictures. Tell me which one has a different ending sound.	seal, cat, pail
Oddity tasks: Mid	Listen to the names of these pictures. Tell me which one has a different middle sound.	jack, cap, run
Blend body-coda	I am going to say a word in two parts. When you have heard both parts, you need to say what the whole word is.	/tu/ g
Blend onset-rimes	I am going to say a word in two parts. When you have heard both parts, you need to say what the whole word is.	/w/ eek
Blend phonemes	I am going to say a word in parts. When you have heard all the parts of the word, you need to say what the whole word is.	/s/ /a/ /ve/
Segment onset-rimes	Separate the word by saying the first sound and then the rest of the word:	boat
Segment phonemes	Say each sound you hear in the word	job
Phoneme deletion	Listen to the word ___. Take away the first sound, what is left?	Listen to the word book. Take away the /b/ sound, what is left?
Phoneme substitution: Beginning sounds	If I say the word man and change the first sound to /p/, the new word is pan.	Change the first sound in cat to /h/. What is the new word?
Phoneme substitution: End sounds	If I say the word rat and change the last sound to /g/, the new word is rag.	Change the last sound in can to /p/. What's the new word?
Phoneme substitution: Mid sounds	If I say the word pan, change the middle sound to /i/, the new word is pin	Change the middle sound in cat to /o/. What's the new word?

Crossing linguistic boundaries: Making the most of cross-linguistic influence in the language classroom

Brittney O'Neill, Josh Bennett, & Chantal Vanier
University of Victoria
bkoneill@uvic.ca

Additional language acquisition is often impacted by learners' previous language experience. Though this interaction is often assumed, instructors and learners may not be aware of the actual extent of the effects of cross linguistic influences. This paper offers a general survey of areas in which cross-linguistic influence facilitates or inhibits additional language learning. With increased awareness of potential cross-linguistic influences, both instructors and learners can improve the learning experience by taking advantage of areas of facilitation and paying greater attention to managing inhibitive influences of previous language experience.

Keywords: cross-linguistic influence, language acquisition, language transfer, instruction

1 Introduction

All languages are different; each has its own way of expressing thoughts, desires, experiences, and needs. Cross-linguistic influence[1] is the expression of these differences in individuals who are trying to learn additional languages. Instructors must be aware of all the ways that other languages can influence the language which they are attempting to teach and how they may effectively address negative influence and take advantage of any positive transfer. In this paper, we will first discuss the negative aspects of cross-linguistic influence, followed by the positive, and lastly offer suggestions of how to most effectively manage influence of both types.

[1] Cross-linguistic influence is closely related to language transfer: the process by which speakers use forms and systems from one language in the production or interpretation of another language. This process may occur bi-directionally between languages.

2 Impacts of cross-linguistic influence

2.1 Negative impacts of cross-linguistic influence

Negative cross-linguistic influence has been researched extensively as an offshoot of the once very popular comparative linguistics field (Aslin et al., 1981; Best, McRoberts & Sithole, 1988; Boroditsky, 2000; Cutler et al., 1992; Juffs, 1998; Pytlyk, 2008; Streeter, 1975; Trehub, 1976; Tsushima et al., 1994; Werker et al., 1981; Yip & Matthews, 2000). The influence of a previously learned language on a target language is notable in every aspect of learning, from phonetics and prosody, to morpho-syntax, to semantic classification, to genre-specific styles and idea organization. Though one cannot claim that all language learning difficulties are a result of cross-linguistic influence, this section discusses the negative implications of this influence on sound perception and production, morpho-syntax, semantic interpretation, and even written organization.

2.1.1 Sound perception

Hearing a language being spoken is an important aspect of additional language learning but, due to previously learned languages, some may struggle to perceive what is being uttered. Much research has been carried out on the universal sound distinction abilities of babies, showing that infants can distinguish not only phonemes of their own language but those of others, which are often significantly different (Aslin et al., 1981; Best, McRoberts & Sithole, 1988; Streeter, 1975; Trehub, 1976; Tsushima et al., 1994; Werker et al., 1981; Werker & Tees, 1984). After an individual reaches about one year of age, however, he/she loses this amazing capacity (Werker & Tees, 1984) and becomes a categorical listener (Liberman, Harris, Hoffman & Griffith, 1957). As first demonstrated by Liberman et al. (1957), sounds lie on continua, but once an individual reaches the age of 11–13 months old, he/she ceases to hear along a continuum and instead hears all speech sounds as members of distinct categories. Kuhl (2000) also showed that the phonetic inventory of a speaker's first language has a lasting effect on the organization of his/her auditory discrimination system. Miyawaki et al. (1975) found that adult Japanese speakers were unable to differentiate English phonemes /l/ and /r/ above the level of chance. The ability to differentiate sound does, however, remain so long as sounds are presented in rapid succession (Van Hessen & Schouten, 1992). This suggests that categorical perception is language specific; one still has the auditory ability and need only foster it so that it can help to redefine the linguistic sound categories of the first language.

Segmentation is also vital to processing the speech that we hear. Languages that are timed differently (stress based like English, versus syllable based like French or Mandarin), are also segmented differently. Unfortunately, listeners seem to be monolingual in their segmentation patterns even if they are bilingual speakers (Cutler et al., 1992). According to Cutler et al. (1992), French speakers treat all input whether French or otherwise as input that can be segmented syllabically. When such segmentation fails, the incoming word must be reanalyzed. Assuming that this holds for language users in general, learners seem unable to learn new ways of segmenting speech on an immediate, automatic level.

2.1.2 Sound production

If a learner is unable to hear a sound as unique, they will likely face difficulty producing it. But, even if students can hear a sound, they may assimilate it to a sound from their own language that is similar to the target but easier for them to produce, thereby bringing about foreign accents. Flege et al. (1997) found that speakers were able to produce more intelligible English vowels if similar contrasting vowels existed in their native vowel inventory. Accents are not to be deemed problematic but students may be able to reach a more native-like quality of phoneme production by ways of phonetically based instruction.

Speech melodies are also important to fluent production and are heavily influenced by one's native tongue. Newborn babies produce cry melodies that mimic the speech melodies of their parents (Mampe et al., 2009). From birth, humans seem to be predisposed to a certain pattern of intonation. Pytlyk (2008) discovered a similar phenomenon with English speaking learners of Mandarin. The Mandarin particle *ma*, when added to the end of a statement, turns it into a yes/no question and should be produced in a neutral (flat and relatively low) tone. Pytlyk (2008) found that English speakers studying Mandarin, however, consistently give *ma* a rising tone to mimic the rising intonation pattern of an English question.

Also of interest are the issues of stress versus syllable timing. Varieties such as Singlish, which is syllable timed (Low, Grabe & Nolan, 2000; Deterding, 2001), are often characterised by their non-standard timing or "staccato effect" (Brown, 1988). This rhythmic similarity to languages such as Mandarin, which are also spoken commonly in Singapore, may suggest a certain level of linguistic interference. The natural stress timing of English has shifted to more syllable timing, perhaps in response to the linguistic backgrounds of the speakers in Singapore.

2.1.3 Morpho-syntax

The way that languages encode grammatical and lexical information also differs greatly. A language may be isolating: one word corresponds to one morpheme; inflectional: one word corresponds to one lexeme and various grammatical affixes; agglutinating: one word corresponds to several lexemes (up to one phrase) and various grammatical affixes; or polysynthetic: one word corresponds to multiple lexemes (up to a sentence) and all necessary grammatical morphemes. (Halvor Eifring & Rolf Theil, 2005)

In his work with Japanese, Chinese, and Romance speaking learners of English, Juffs (1998) found that speakers of languages that must encode causation in verbal events overtly (Japanese and Chinese) have significantly more trouble with ambiguous English sentences involving causative verbs than do speakers of Romance languages, which use a similar causative verb encoding system to English. This sort of interference can also occur in native bilinguals as shown by Sanchez (2006). Bilingual Kechwa–Spanish children were found to use a non-traditional Spanish structure, not used by speakers of Spanish alone, which reflects their use of desiderative affixes to convey volition in Kechwa. In this instance, there is an important interaction between a speaker of two native languages leading to atypical uses. Hence, teachers of students working from a language with different morphological patterns than the target language are likely to encounter unusual forms of language production, as learners attempt to assimilate the new language to the structure of the old.

Even in languages with similar uses of inflection such as Spanish and Italian, word order and other purely syntactic processes can complicate the acquisition process. Argument and predicate ordering, the manner in which phrases are placed in a sentence, i.e., how the subject, verb, object, and indirect object are organized, is a particularly important aspect of syntactic variation as languages may organize sentences as SVO, SOV, or OSV. Another important aspect to consider (and to remind students of) is the positioning of adjectival phrases. Languages like French tend towards placing adjectival structures after the noun that they modify, whilst languages like Mandarin place all modifiers before the noun. Languages like English, however, tend to place adjectives before nouns but adjectival phrases (i.e., ‘that’ or ‘which’ phrases) after the noun (Huang, 2010).

As seen in Yip & Matthews’s (2000) study of a bilingual Cantonese–English speaking child, transfer of word ordering is present in language use. The child showed interference with structures such as Wh-questions but only from Cantonese to English. The child was believed to be equally proficient in both languages and so the prominence of Cantonese structures such as the Wh-in-situ-questions, which retain the question word where the answer would be in a declarative sentence, suggests markedness, as detailed by Eckman (1977, 1981,

2004)'s Marked Differential Hypothesis,[2] as a determinate in cross-linguistic influence.

2.1.4 Semantic variation

Morpho-syntactic interference may require overcoming deep habits, but semantic variation may demand an entirely new conception of the material at hand. The Sapir-Whorf hypothesis proposes that this difference in linguistic conceptualization actually constrains the way that we perceive the world (Whorf, 1956). The strong version of the hypothesis has (for the most part) been discarded, but the weak version (suggesting that the way our language partitions the world influences the way we perceive the world) remains of interest (Gentner & Goldin-Meadow, 2003).

Boroditsky (2000), examined Mandarin time words and showed that Mandarin speakers can be primed differently than English speakers in the domain of time, suggesting some influence of language on thought as is entailed by the weak Whorfian hypothesis. For example, Mandarin uses vertical metaphors (e.g. the past (*earlier*) is *up* or *shang* and the future (*later* is *down* or *xia*) for time as well as horizontal metaphors (e.g. *before* can be expressed by *in front of* or *yi qian*) analogous to English. This sort of variation in semantic content of expressions can be very troubling for students of additional languages.

Number systems also differ from language to language both in terms of base and encoding style used. English is a good example of a marked base ten number system that has non-transparent terms for numbers like *thirty-four*. Mandarin on the other hand encodes numbers in an extremely transparent manner: *thirty four* is *san-shi-si* or *three-ten-four*. Mandarin speaking children develop numeracy skills faster than English children (Anuio et al., 2009) which may suggest that the way our language encodes numbers may influence our general conception of number systems. Thus, for learners of languages that use less transparent encoding systems, the language specific preferences in encoding must be kept in mind and explicitly taught.

Languages differ, not only in conceptualization but also in categorization. One of the most studied domains is that of colour. Heider (1972) claims that there is no deep effect of language on colour perception or memory in English and Dani speakers, but these findings have been questioned as Dani colours overlap,

[2] In this hypothesis, Eckman suggests that all features of language are more or less marked in accordance with their frequency of use across languages, ease of acquisition and complexity. For example: a syllable beginning with a consonant cluster is more marked than a syllable beginning with a single consonant. Eckman also argues that language structures are learned in order of markedness and are more or less difficult to learn in accordance with their markedness (Eckman, 1977, 1981, 2004).

but do not crosscut, English colour boundaries. Roberson et al. (2000), working with Berinmo, a language which regularly crosscuts English colour classes, did find a difference in recognition and memory using similar methodologies to Heider (1972). Lastly, Winawer et al. (2007) found significant differences in colour recognition in Russian and English speakers with blues (in Russian there are two separate terms, one for light blues one for dark). Winawer et al. (2007) also found that, when a linguistic interference task (e.g. reading a string of nonsense syllables aloud whilst doing the recognition task) was used, both groups performed equally poorly. This may suggest that perception remains unchanged but is filtered through linguistic categories: a difference in conception, not perception.

2.1.5 Organizational habits

A final area of cross-linguistic interference lies in how learners and speakers of a language organize ideas and present thoughts. As Huang (2010) states, languages differ in their preference of topic or subject prominence, extent of use of connecting forms, and choice of ordering information. An instructor must be aware of the presence of these habits and their potential to create miscommunications, stylistic problems, and ambiguities in the written work of students.

2.2 Positive impacts of cross-linguistic influence

There has been little research done, within the field of second language acquisition, on the positive or potentially positive effects of knowing a previous language on learning a new language. Thus, we discuss some potential positive influences which deserve consideration.

Language learning can be aided when the target language is in the same language family and/or shares linguistic roots with a language already known to the learner. Cognates, words that are similar or the same between two languages, such as *night* (English), *nuit* (French), *nacht* (German), *natt* (Swedish, Norwegian), *nótt* (Icelandic) as well as *apple* (English) and *appel* (Dutch), will in many cases require little to no effort for the learner to acquire. These cases can also make it easier for the learner to puzzle out the meaning of an unfamiliar word in the target language, due to common linguistic roots of words in the target language and previously known language. Coming to the meaning of the word in this way can not only aid the learner in recollection of the word, but also give them greater confidence in learning and speaking the language, leading to better overall language learning. It can also be easier for a learner when they are already

familiar with the sentence structure of a target language, such as a subject-verb-object (SVO) language speaker learning another SVO language, as a new sentence order need not be learned.

Also to consider is the effect that previous language learning can have on further attempts to learn additional languages. Many methods of language learning give the learner increased knowledge of linguistic structures and principles, and having this knowledge already available can potentially make the language learning process easier. For example, an English speaker who has already learned French should have some knowledge of the inflected imperfect tense[3] (which normal monolingual speakers of English do not). This makes using the tense in another language that uses it (such as Latin) far easier. Some learners will benefit from being able to sort their new knowledge into formal terms and categories, such as subject-verb-object sentence order, inflection, affixation, and other formal grammatical terms as it provides a skeleton onto which the flesh of the target language can be associated and connected.

Although these examples of positive cross-linguistic influence are not necessarily substantiated by research, at the very least they indicate the potential for study in this particular area of language learning.

3 Teaching implications

Without turning the classroom into a comparative linguistics lecture, language instructors may find it useful to know as many contrasts and common aspects between the language that they are teaching and the languages used by students as possible. Instructors cannot help learners to be aware of potential biases (and how to avoid them) if they are not aware of the biases themselves. This section then will outline some helpful ideas for using knowledge of cross-linguistic influences to help learners within a language learning environment.

3.1 Phonetically-based instruction

We believe that language learning and intelligibility of speech can be greatly improved by a phonetic approach to sound production. Though some learners may be able to produce sounds by listening alone, many also need a little extra help. In teaching non-English phonemes/phoneme contrasts verbal descriptions of mouth/tongue/lip alignment and even small diagrams may help simplify a challenging task. If taught how to position their tongue and lips, students are

[3] Usually considered analogous to the English construction *was* VERB*ing*, it indicates a past continuous or incomplete action.

more likely to be able to produce the sound than if merely listening to an instructor and guessing. For example, the first author encountered an English speaking learner of Mandarin who had, for the entire two and a half years he had been studying the language, believed himself unable to produced the difference between /tʂʰ/ and /tɕʰ/ (having pronounced both as the English /tʃ/). He was able to acquire a noticeable improvement in pronunciation in less than five minutes of phonetically guided practice with the guidance of a rough picture of tongue positions sketched out by the instructor.

For example, if one wants to produce the /θ/ found in English, it seems to be more helpful to offer a detailed description of how the sound is produced than to just produce it in an attempt to elicit a response. It is also often helpful to employ analogies to known sounds. For a speaker of Japanese attempting to produce the /θ/, an instructor may wish to have them first produce an [s] sound. Then either description or drawing can show the student how their tongue is placed in the mouth near the alveolar ridge (the hard part of the roof of the mouth just behind the teeth) for the [s]. Now the student is asked to try to make a [s] sound while gradually extending the tongue towards and between their teeth, they will likely have far greater success. Many of the authors' undergraduate peers of various departmental affiliations have been recently subjected to crash courses in non-English phonemes. Even for those who were unable to hear the contrasts when the sounds were made, a careful description of what their tongue needed to do yielded (without fail – though with a multitude of complaints about unnaturalness of the sound) the phoneme in question. Instructors may also find it useful to consult the forthcoming Truong (2010), which explores phonetic pronunciation instruction among learners of L2 Japanese.

3.2 Cognates: (False) friends

There are many words across languages that, though similar in spelling and pronunciation, do not have the same meaning. For example, *bekommen*, in German, does not mean *to become* but, instead, *to get/acquire*. There are also many words which can provide useful handholds for learners of an additional language. Even something as simple as *appel* and *apple* (see section 2.2) can, if nothing else, allow the learner to feel more at ease with a language.

Instructors, however, must make their students aware of those cognates which are not in fact synonymous in meaning as learners may be tempted to over-extend analogies to their previously learned languages. As discussed in Prator (1967), though acquiring nearly identical tokens across languages is the lowest of difficulty, acquiring a token which, on the surface, is very similar but is actually connected to a different function of meaning is amongst the most difficult to learn. Thus, false friends deserve exposition in the language classroom.

3.3 Imagery

When facing semantic differences in language's conceptualizations or categorizations of the world, instructors face a unique challenge. Asking students to imagine the conceptual domain at hand in a clear visual image that aligns with the target language can offer a much more salient understanding of the system. Drawing pictures of the idea can also be helpful. Learners of Mandarin, for example, may wish to imagine that one is lying on their back in a field with time rising around them so that they see the past immediately after it happens but cannot see the future which is below them, hence the past is *up* or *shang* and the future is *down* or *xia* – as discussed in section 2.1.4. The first author has found this particular practice immensely helpful in her ongoing attempt to learn Mandarin.

Visuals can also be used in domains such as colour. For example, in languages that have differing colour systems to English, such as Russian with its two blues mentioned previously, students may be assisted by actual colour chips showing the prototypical shade for a particular colour name, or even a whole hue chart with each colour name marked off in its own unique section. Though not all students are visual learners, visualization can be a good place to start in giving learners access to varied conceptual systems from which they may be able to expand the mnemonics to suit their specific learning styles (e.g. Auditory learners may be receptive to rhyming or acrostic based oral mnemonics).

3.4 Finding patterns and remaining aware

Instructors continually try to find ways to help students learn; knowing their previous language background can be immensely helpful in this area. As we have discussed previously, languages do differ, and so instructors may be able to improve students' learning by finding ways to tap into their existing knowledge of linguistic forms and the commonalities between those forms and the forms used by the target language.

Instructors must also remain aware of the differences that can lead to difficulty. It cannot be denied that linguistic transfer occurs as students grapple with a new language, often using an existing language as a medium. It is important to be knowledgeable about cross-linguistic similarities and differences and to offer students side by side comparison and exposition of salient features of the languages at hand. For example, if studying past tense systems in English, it may be useful to discuss how past tense systems work in learners' previously learned languages for comparison. Students can improve their learning by seeing the differences and similarities and finding patterns that allow them to avoid inter-language errors. Students aware of the patterns in their existing languages

and encouraged to recognize the patterns of the target language are more likely to find ease in the process of reanalysis of words and structures that may be difficult to segment meaningfully.

4 Conclusion

In order to take full advantage of working with and around cross-linguistic influence in the classroom, more research is necessary on positive cross-linguistic influence to add to the considerable amount of research done concerning negative cross-linguistic influence. Instructors should be fully aware of the deeper linguistic characteristics of both the target language and the languages that the students have previously learned in order to maximize teaching potential. Teaching a language is not always about breaking the habits of the previous language or languages, it is also about becoming proficient in an additional communication medium. As such, instructors require many skills and insights above and beyond knowledge of cross-linguistic influence. That said, making the most of the impacts of contrasts and commonalities between previously learned and target languages constitutes an important part of language learning and teaching as a whole.

References

Aslin, R.N., Pisoni, D.B., Hennessy, B.L. & Perey, A.J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development* 52, 1135–1145.

Aunio, P., Aubrey, C., Godfrey, R., Yuejuan, P. & Liu, Y. (2008). Children's early numeracy in England, Finland and People's Republic of China. *International Journal of Early Years Education* 16(3), 203–321.

Best, C.T., McRoberts, G.W. & Sithole, N.N. (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 345–360.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43, 1–22.

Brown, Adam (1988). The staccato effect in the pronunciation of English in Malaysia and Singapore. In: J. Foley (Ed.), *New Englishes: The case of Singapore*. Singapore: Singapore University Press, 115–128.

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24(3), 381–410.

Deterding, D. (2001). The measurement of rhythm: A comparison of Singapore and British English. *Journal of Phonetics* 29, 217–230.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning* 27, 315–330.

Eckman, F. (1981). Markedness and degree of difficulty in second language acquisition. In: J.G. Savard & L. LaForge (Eds.), *Proceedings of the fifth Congress of the International Association of Applied Linguistics*. Laval, PQ: Les Presses de l'Universite Laval, 115–126.

Eckman, F. (2004). Optimality Theory, markedness, and second language syntax: The case of resumptive pronouns in relative clauses. *Studies in Phonetics, Phonology, Morphology* 10, 89–110.

Eifring, H. & Theil, R. (2005). Linguistics for students of Asian and African languages. Unpublished manuscript.

Flege, J.E., Bohn, O. & Jang, S. (1995). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25, 437–470.

Gentner, D. & Goldin-Meadow, S. (2003). Whither Whorf. In: D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind.* Cambridge, MA: MIT Press, 3–14.

Heider, E.R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology* 93, 10–20.

Huang, L. (2010). The potential influence of L1 (Chinese) on L2 (English) communication. *ELT Journal* 64(2), 155–164.

Juffs, A. (1998). Some effects of first language argument structure and morphosyntax on second language sentence processing. *Second Language Research* 14, 406–424.

Kuhl, P.K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences* 97, 11850–11857.

Liberman, A.M., Harris, K.S., Hoffman, H.S. & Griffiths, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 53, 358–368.

Ling, L.E., Grabe, E. & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Journal of Phonetics* 29, 217–230.

Mampe, B., Friederici, A.D., Christophe, A. & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology* 19(23), 1994–1997.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Jenkins, J.J. & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18(5), 331–340.

Prator, C. (1967). Hierarchy of difficulty. Unpublished classroom lecture, University of California, Los Angeles.

Pytlyk, C. (2008). Interlanguage prosody: Native English speakers' production of Mandarin yes-no questions. *Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association.*

Roberson, D., Davidoff, J., Davies, I. R., Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology* 50(4), 378–411.

Sanchez, L. (2009). Kechwa and Spanish bilingual grammars: Testing hypotheses on functional interference and convergence. *International Journal of Bilingual Education and Bilingualism* 9(5), 535–556.

Schouten, M.E.H. & Van Hessen, A.J. (1992). Modelling phoneme perception, I:

Categorical perception. *Journal of the Acoustical Society of America* 92, 1841–1855.

Streeter, L.A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259, 39–41.

Trehub, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47, 466–472.

Truong, A. (2010). *Linguistically based pronunciation instruction in L2 Japanese*. MA thesis, University of Victoria.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Komo, M., Menyuk, P. & Best, C. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants 6–12 months: Language-specific developmental changes in speech perception abilities. Paper presented at the International Conference on Spoken Language. Yokohama, Japan.

Werker, J.F., Gilbert, J.H.V., Humphrey, K. & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development* 52, 349–353.

Werker, J.F. & Tees, R.C. (1984). Crosslanguage speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49–63.

Winawer, J., Witthoft, N., Frank, M.C., Wu, L., Wade, A.R. & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Science U. S. A.* 104(19), 7780–7785.

Whorf, B. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Yip, V., Matthews, S. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition* 3(3), 193–208.

The effects of time compression on the comprehension of natural and synthetic speech

Janine Lebeter & Susan Saunders
University of Victoria
jlebeter@uvic.ca

Synthetic speech is commonly used as the output signal in text-to-speech synthesizers. The purpose of this study is to determine if high quality synthetic speech, such as the type used by speech-generating devices, is perceived as well as natural human speech. Little research has looked at the comprehension of synthetic versus natural speech through the dimension of time compression. This research fills that gap by comparing the comprehension of time-compressed natural speech signals and time-compressed synthetic speech signals. A secondary aim is to determine the quality of current text-to-speech (TTS) synthesizers that come with current (as of 2010) computers. In this experiment, signal comprehension was tested with a speeded sentence verification task. Participants were able to verify natural speech sentences faster and more accurately than synthetic speech sentences. Additionally, as the sentence compression rate was increased, comprehension became more difficult for both speech conditions, with the greatest adverse affect being found for synthetic speech comprehension.

1 Introduction

Synthetic speech generators have become an important tool in the lives of many individuals. It is common for people with language disorders and delays to use speech synthesizers to augment their communication, and within the past few years, both Windows and Macintosh computers have equipped their newest models with pre-installed text-to-speech generators. The majority of these devices use text-to-speech synthesis wherein graphemes, digits, and words are entered using a keyboard or touch screen as input, which is then converted into a synthetic speech waveform by a set of algorithmic rules (Koul, 2003; Koul & Clappsaddle, 2006). Studies investigating what effect, if any, speech generating devices (SGDs) have on the lives of individuals with mild to severe intellectual disabilities, visual impairments, and special communication needs have clearly shown that SGDs make a profound difference. These devices have been shown to be more effective than vocalizations, gestures and non-electronic communication

boards in conveying information, and they lead to an increase in positive communicative interactions with peers and support personnel (Koul, 2003; Koul & Clapsaddle, 2006; Koul & Hester, 2006). Even though SGDs play such an important role in the lives of many individuals, the quality of the synthetic speech is not guaranteed, and depending on the sophistication of the device, the output may vary greatly (Koul, 2003).

Since the middle to late 1980s, many researchers have compared the quality of synthetic versus natural speech (Hoover et al. 1987; Logan et al., 1989; Mitchel & Atkins, 1989; as cited in Koul, 2003; Mirenda & Beukelman, 1987). Speech signal quality is discussed in terms of intelligibility and comprehension, where intelligibility refers to an individual's ability to recognize phonemes and words presented in isolation, while comprehension requires that a listener transform the linguistic message into a meaningful mental representation (Koul, 2003). The present study tests for signal comprehension because it is important that people are able to construct meaningful mental representations from the synthetic speech used in speech generating devices. Koul (2003) suggests that, for single word identification tasks conducted in ideal listening conditions, there is no significant difference between the perception of high-quality synthetic speech and natural speech. Other research has supported the opposite view: that digitized or synthetic speech is more difficult to perceive than natural speech and demands greater cognitive resources to process (Duffy & Pisoni, 1992; Francis & Nusbaum, 2009; Mirenda & Beukelman, 1987). Since the advent of the first text-to-speech computer-based system in 1968, formant synthesis technology has greatly improved. One aim of the present study is to assess the quality of current text-to-speech synthesizers that come pre-installed in new computers.

Past research has tested the quality of synthetic speech by manipulating variables such as background noise, age of listener, intellectual ability of listener, and experience with the signal (Koul, 2003; Koul & Hanners, 1997; Mirenda & Beukelman, 1987). However, to the best of our knowledge, no study has tested the quality of synthetic speech by manipulating speech rate. For participants, fast speech rates create an adverse listening condition (Adank & Devlin, 2010; Adank & Janse, 2009; Dupoux & Green, 1997; Golomb, Peelle, & Wingfield, 2007; Janse, 2004; Pallier & Sebastian-Gallés, 1998), which is desirable when signal quality is being tested. Additionally, Dupoux and Green (1997) have pointed to time-compressed speech as being an ideal independent variable for a number of reasons. Firstly, with linear time compression, speech signals can be altered in quantifiable and measurable ways to create stimuli that are outside the bounds of everyday experience. Secondly, newer compression algorithms such as Praat's "Pitch Synchronous Overlap and Add" (PSOLA) function (Boersma & Weenik, 2009), used in the present study, allow speech to be compressed without deleting segments of the original signal or creating discontinuities, which was common with older compression techniques. Finally, compressed speech affects the

perceived rate at which the signal was produced. Because of this, it has been argued that any perceptual effects found for time compressed speech can be compared with, and generalized to, more natural changes in speech rate (Dupoux & Green, 1997).

Adank and Janse's (2009) study of perceptual learning mechanisms used naturally fast and linearly time-compressed speech to study human adaptation to atypical speech signals. Participants were asked to perform a speeded sentence verification task for both naturally fast and artificially time-compressed stimuli. Surprisingly, the researchers found that time-compressed speech was easier to adapt to – as measured by faster reaction times and overall higher verification accuracy – than natural fast speech. This finding supports past research, which had found that natural fast speech is difficult to adapt to because it is not only temporally compressed, but it is also spectrally different from regular conversational speech (Janse, 2004; Adank & Janse, 2009). The spectral variation that occurs with fast speaking rates is caused by the increased occurrence of coarticulation and segment deletion, a change in the overall prosodic pattern of the speech stream, and a tendency to reduce the duration of vowels and unstressed syllables.

The naturally fast stimuli used in Adank and Janse's (2009) study were all produced by a single individual. The speaker was instructed to read 180 experimental sentences aloud as declarative statements at his normal speaking rate, while recordings were taken. He was then instructed to produce all of the stimuli again by reading each individual sentence aloud four times in quick succession so as to achieve a very fluent speaking rate. It was found that, on average, the fast versions of the sentences were compressed to approximately 46% of the original sentence duration, with the fastest items being produced at approximately 33% of the original sentence duration (Adank & Janse, 2009). Given that such fast speech rates are achievable by human articulation, we predicted that our participants would be able to comprehend at least some sentences presented at such fast rates, as they will have had experience with these fast speech rates during their lifetime. Dupoux and Green (1997) also analyzed perceptual adjustment mechanisms for highly compressed natural speech. Their fast stimuli were compressed to 38% and 45% of the original speaking rate. It was found that the sentences compressed to 38% of their original duration were difficult for participants to understand, and that the adjustment process took longer for stimuli that had been compressed to a greater degree (Dupoux & Green, 1997). The literature shows that increased speech rates are more difficult to perceive and comprehend than conversational speech rates (Adank & Devlin, 2010; Adank & Janse, 2009; Dupoux & Green, 1997; Golomb, Peelle, & Wingfield, 2007; Janse, 2004; Pallier & Sebastian-Gallés, 1998). For the present study, it is predicted that, as the compression rate is increased across experimental blocks, signal comprehension will become more difficult in both

the natural and synthetic stimuli conditions. The speech signal that facilitates comprehension at a higher rate of compression will be considered to be the higher quality signal because it allows for comprehension in the more difficult listening condition. We predict that for a more difficult task such as the speeded sentence verification task employed here, participants will not comprehend the synthetic speech as well as they comprehend natural speech. Furthermore, it is predicted that an increased sentence compression rate will have a more negative effect on synthetic speech perception than on natural speech perception.

The overall findings in previous literature can be summarized as follows. a) Naturally spoken words and sentences have typically been shown to be more intelligible and comprehensible than synthetic speech; however, depending on the sophistication of the synthesizing device there may be no noticeable difference in speech quality. b) Time-compressed speech is preferred over naturally fast speech, and c) fast speech rates are more difficult to comprehend than normal, conversational speech rates.

While past studies have compared fast natural speech with time-compressed natural speech, there is a research gap with respect to the comparison of time-compressed natural speech with time-compressed synthetic speech. The aim of this study is to address this gap in the literature. Specifically, by manipulating the variable of speech rate, we determine whether synthetic speech is comprehended as well as natural human speech. Secondly, we establish whether or not comprehension deteriorates equally for both speech signals as the speech rate increases. Based on these results, we assess the quality of current (2010) text-to-speech generators that come pre-installed with Windows and Macintosh computers. Since much of the research on synthesized speech took place over 20 years ago, we predict that the quality of synthesized speech will have improved. If the sophistication of text-to-speech generators has significantly improved, we predict that the participants who are presented with synthetic time-compressed speech will not perform significantly better[1] or worse than those who are presented with natural time-compressed speech. Conversely, if the quality of speech synthesizers has not improved over the last two decades, we predict that the participants who are presented with synthetic time-compressed speech will perform worse than those who are presented with natural time-compressed speech.

[1] "Better" is quantified as a faster reaction time and higher percent accuracy for the speeded sentence verification task.

2 Materials and methods

2.1 Participants

Twenty-five participants (12 male, 13 female) took part in the study. All participants were native Canadian English speakers between the ages of 18 and 30. They reported having limited linguistics training, no major previous exposure to time-compressed or synthesized speech and no hearing loss, although no audiometric test was given. All participants gave their written informed consent to participate in the study, and were not paid or compensated for their time.

2.2 Speech stimuli

This experiment included two sets of auditory stimuli: one synthetic speech set and one natural speech set. Each stimuli set contained recordings of the same 96 true-or-false sentences adapted from Baddeley, Emslie, and Nimmo-Smith's (1992) Speech and Capacity of Language Processing Test, or SCOLP, which was used by both Adank and Janse (2009) and Adank and Devlin (2010). The sentences were slightly altered from their original format: new subjects and predicates were substituted for the original content words. The substituted lexical items were all common, high frequency English words. Only high frequency English words were used in order to avoid a possible confound stemming from lexical confusion. Although the sentential content was altered, the general format of the SCOLP sentences was preserved because SCOLP sentences have been widely tested and have proven to be a reliable measure of language comprehension (Adank & Janse, 2009). A complete list of the sentence stimuli used in the present study is given in the Appendix.

The statements made in the sentences were all unambiguously true or false (e.g., "An ant is a small insect." versus "Elephants are small insects.") in order to ensure that each statement was verifiable. Each true sentence had a false sentence counterpart, as demonstrated in the above example, thus 48 pairs of sentences were used in the experiment. All of the sentences were 7 or 8 syllables long, in order to avoid a possible confound of variable sentence length. A number of past studies have used syllables as the unit of measurement when controlling for sentence length or for the speed of sentence presentation (Adank, & Devlin, 2010; Adank & Janse, 2009; Dupoux & Green, 1997; Janse, 2004).

Of the 96 sentences, 16 were used for pre-test training. The remaining 80 sentences were divided into 5 experimental blocks, with each block containing 16 trial sentences, as is shown below in Table 1. These 16 trials were semi-randomized within their respective blocks. There were an equal number of true and false statements within each block, and sentence pairs were distributed across

blocks so pairs did not occur together. The sentences were linearly time-compressed to five different percentages of their original duration using Praat's Lengthen **(**Add-Overlap) function under "Synthesize > Convert" (Boersma and Weenik, 2009). The compression rates used were: 42%, 40%, 38%, 36%, and 34% of the original sentence durations.

Table 1. Experimental Design. Each block contained 16 sentences, followed by 3000 m.s. to respond. Once the answer was recorded, there was a 100 m.s. silence before the next sentence began. Compression rates (%) ranged from 44% to 34%.

Practice (44%)	Block 1 (42%)	Block 2 (40%)	Block 3 (38%)	Block 4 (36%)	Block 5 (34%)

These compression rates were selected on the basis of past research and a participant pre-test. It was decided that each block should include 16 trial sentences because past research has demonstrated that comprehension of a rapid or unusual signal improves over time, and that normalization typically stabilizes after 14–18 sentences worth of experience with a given signal (Adank & Devlin, 2010). The researchers wanted to allow participants a sufficient number of trials at each compression rate so that participants could reach near optimal comprehension performance.

Because the aim of the present study is to compare the comprehension of time-compressed synthesized speech with the comprehension of time-compressed normal speech, two versions of the same 96 sentences were created, one normal speech version and one synthesized speech version. A monolingual female speaker of English from Summerland, British Columbia, Canada recorded the natural speech stimuli. Her recordings were made in a sound-attenuated booth using an M-Audio Luna microphone from the large diaphragm condenser family. The synthesized versions of the experimental stimuli were generated using the text-to-speech "Anna" (Microsoft Inc.) voice that comes included with Windows 7- and Windows Vista-equipped computers. These synthesized sentences were externally recorded with an M-Audio Microtrack solid state recorder. Both the natural and synthetic sentences were clipped to have zero seconds of silence before and after the utterance and saved into 192 separate files. The files recorded by "Anna" were time-compressed or enhanced to be equal length to their natural spoken counterpart.

2.3 Procedure

The study was conducted in the University of Victoria Phonetics Laboratory. All participants received oral instructions read from a script before the experiment began. Participants were randomly assigned to either the synthetic or the natural speech condition. Group A heard natural stimuli, while group B heard synthetic stimuli. The tasks for each group were the same, as were the sentences in each set. This is in accordance with the atypical block-design taken directly from Adank and Devlin (2010). The experiment was run on the software program E-Prime (Schneider et al., 2002a, 2002b). Participants heard the sentence stimuli through headphones at a comfortable sound level, which they determined.

As previously mentioned, the current study will replicate the atypical block design employed by Adank and Devlin (2010). An atypical block design requires that each participant be tested with only one of the two possible signal types. This design is necessary because it has been shown that continually alternating signal type limits behavioral adaptation, thus preventing participants from reaching their optimal performance level (Adank & Devlin, 2010). Because the goal of this experiment is to test the upper limits of synthetic and natural fast speech comprehension, any inhibition of adaptation would be detrimental to the study. Thus, participants were tested on the normal or the synthesized speech signal only.

The participants were first presented with 16 familiarization sentences. The task of the participant was to decide on the validity of each sentence statement presented, and indicate their true-or-false response as quickly as they could with a keyboard button press. Reaction times longer than 3000 m.s. were coded as ‘no response’ and E-Prime automatically presented the next sentence token in the sequence. Both accuracy and reaction time measurements were recorded for each sentence trial. Both measurements were recorded in order to capture in greater detail the cognitive processing costs required for comprehending synthesized and normal speech signals at different compression rates. Reaction time measurements were taken from the end of the audio file following similar previous research procedure (Adank & Janse, 2009; Adank & Delvin, 2010). Good signal comprehension is defined as a high level of response accuracy and short reaction times because these behaviors indicate that the participant was able to easily comprehend and respond to the stimulus presented (Adank & Delvin, 2010).

2.4 Data analysis

Both response accuracy and response time measurements were used as the dependant measures in this study. A total of 25 participants × 96 trial sentences × 2 measurements per trial = 4800 data tokens to analyze (2400 accuracy tokens, and 2400 response time tokens). Accuracy and response time averages were compared between the two speech type conditions and were analyzed across the five compression rate blocks.

3 Results

Table 2 shows the average reaction times and accuracy percentages for the synthetic and natural speech conditions. Overall, participants had shorter reaction times in the normal speech condition versus the synthetic speech condition. The normal speech condition participants also had a higher level of accuracy in their sentence verification responses. Taken together, these two findings suggest that there is a main effect of speech type on comprehension; normal human speech is easier to comprehend than synthetic speech.

Table 2. Average reaction times and percent accuracy across all five blocks.

	Reaction time (m.s.)		Accuracy (%)	
	Normal	Synthetic	Normal	Synthetic
Average:	1015.01	1370.08	85.4	62.5

Figure 1 plots participants' average response accuracy (*y* axis) in making a true or false decision as a function of the signal's compression rate/speed (*x* axis). Average response accuracy for participants in the normal speech condition (black bar graphs), are plotted against the average response accuracy of participants in the synthetic speech condition (grey bar graphs). Figure 1 shows that participants responded more accurately in the normal speech condition than in the synthetic speech condition, for all of the five different compression rates. In the normal speech condition, the lowest response accuracy average was 83.9% and occurred in Block 3 at a 38% compression rate. In the synthetic speech condition, the lowest response accuracy average was 62.5% and occurred in Block 5, at a 34% compression rate. In the normal speech condition, average accuracy rose and fell in random fashion across blocks; there did not appear to be a main effect of compression rate on response accuracy. In the synthetic speech condition average accuracy rose and fell as it did in the natural speech condition, however, there was a general trend that participants in the synthetic speech group became less accurate in their responses as the compression rate was increased.

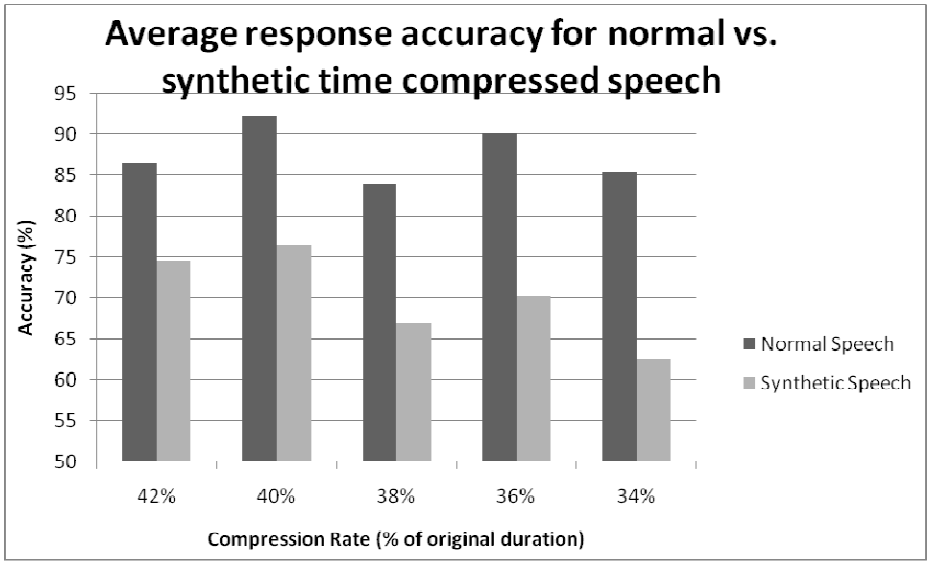


Figure 1. Average response accuracy plotted as a function of speech type and signal speed (compression rate).

Figure 2 plots participants' reaction times (*y* axis) in making a true or false decision as a function of the signal's compression rate/speed (*x* axis). Reaction times for participants in the normal speech condition (black bar graphs) are plotted against the reaction times of participants in the synthetic speech condition (grey bar graphs). Figure 2 suggests that participants in the synthetic speech condition required a longer amount of time to make a sentence verification decision than did the participants of the normal speech condition. When analyzing reaction time performance across compression blocks, we see that in the synthetic speech condition participants' reaction times became steadily slower as the compression rate of the signal increased. In a general way, this effect was also seen in the normal speech condition as well. Figure 2 suggests that there is a main effect of compression rate on decision response time, and that the normal speech signal is easier to perceive than the synthetic speech.

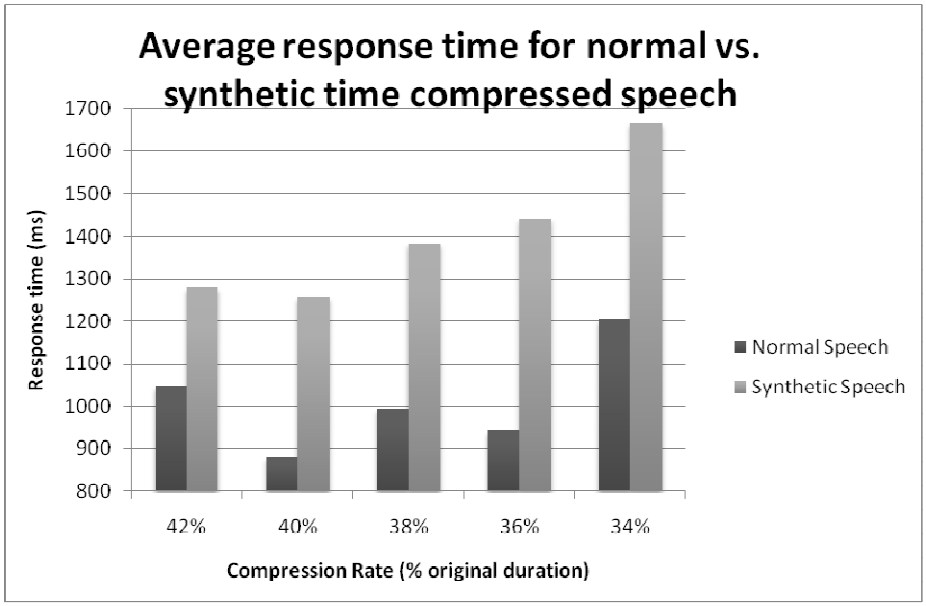


Figure 2. Average reaction time plotted as a function of speech type and signal speed (compression rate).

4 Discussion

4.1 Key findings

The results demonstrate two important points. First, they show that listeners are not able to comprehend synthetic speech as well as they comprehend natural speech. This goes against Koul (2003), who found synthetic speech to be of comparable quality to natural speech, although the majority of the literature supported the view that synthetic speech is more difficult to perceive and requires greater cognitive resources to comprehend (Duffy & Pisoni, 1992; Francis & Nusbaum, 2009; Mirenda & Beukelman, 1987). Listener performance in the present experiment confirms that synthetic speech is more difficult to process than natural speech. Individuals in the synthetic speech condition had longer reaction times and lower response accuracy averages than normal speech participants for all five of the compression rates.

Secondly, the results show that the adverse listening condition of fast speech makes comprehension more difficult for synthetic speech listeners than for natural speech listeners. Both the average accuracy results and the reaction time results support this finding. In both speech conditions, the average accuracy of

participant responses rose and fell as the compression rate was increased. In the normal speech condition, participants' average accuracy for Block 1 (42% compression) was 86.5% and their average accuracy for Block 5 (34% compression) was 85.4%. This indicates that increasing the speed of natural speech does not have a large effect on participants' ability to accurately verify sentences. In the synthetic speech condition, however, increasing the speed of the signal did affect comprehension performance. The average response accuracy for the synthetic speech group in Block 1 was 74.5%, and this already low figure dropped to 62.5% accuracy by Block 5, a difference of -10%. The reasons why synthetic speech comprehension may have been so adversely affected by a fast signal presentation rate will be discussed in detail below.

The reaction time results also suggest that the comprehension of synthetic speech is more affected by an increased signal rate than normal speech. In the normal speech condition, the average reaction time of participants rose and fell across blocks, although the general trend was that reaction times became longer as the presentation rate increased. For the normal speech condition, Block 2 (40% compression) had the shortest average reaction time of 881.00 m.s. Block 5 (34% compression) had the longest average reaction time of 1204.89 m.s. The difference between the fastest and slowest reaction time averages for participants in the natural speech group was +323.89 m.s. The shortest average reaction time in the synthetic speech condition occurred in Block 2 (40% compression) and was 1256.59 m.s.; the longest reaction time average occurred in Block 5 and was 1666.40 m.s., a total difference of +409.81m.s. The fact that there is a larger reaction time difference for the synthetic speech group than for the natural speech group suggests that participants in the synthetic speech condition were more adversely affected by the increase in speech rate.

In sum, our results show that listeners are not able to comprehend synthetic speech as well as they are able to comprehend natural speech, and that an increase in speech rate has a greater adverse affect on synthetic speech perception than on natural speech perception. These findings are in line with our original hypotheses. Despite the technological advances that have greatly improved the quality of synthetic speech in recent years, there are a variety of possible reasons why people are still unable to comprehend synthetic speech as well as they comprehend natural human speech. First, let us consider the Windows 07's "Anna" voice that was used in the present study. The Microsoft "Anna" voice was created with formant synthesis technology. In formant synthesis, the different acoustic parameters of speech such as fundamental frequency, voicing, and signal amplitude, et cetera, are produced by algorithmic rules, which create the artificial speech waveform. For this type of speech synthesis, it is common that only one or two acoustic cues will be specified to distinguish a given phoneme, and often the same acoustic cues are used for more than one phoneme. Researchers Francis and Nusbaum (2009) identify this impoverished and misleading cue structure as

being the primary reason why synthetic speech perception can be so difficult. In natural speech, there are multiple acoustic cues that interact to create the percept of a specific phoneme. In synthetic speech, on the other hand, perceptual ambiguity may be increased because (1) fewer discrete cues have been encoded, so the relationships between the synthesized acoustic cues may be uninformative and misleading in comparison with the cue structure of natural speech, and (2) the same patterns of acoustic cues appear in a greater range of contexts for synthesized speech (Francis & Nusbaum, 2009). This acoustic-phonetic ambiguity, which is found in synthetic speech, is one possible reason why synthetic speech comprehension is difficult. Also, because speech synthesizers are engineered by humans, there is always the possibility that human engineering errors could result in misleading cue structure (Francis & Nusbaum, 2009). In such circumstances, the listeners would need to learn to inhibit their perceptual intuitions for the poorly engineered contexts in question.

Another possible reason why the synthetic speech stimuli may have been more difficult to comprehend is that this study tested for listeners' comprehension of synthetic speech rather than just the intelligibility of the signal itself. Comprehension requires a higher level of cognitive processing than does simple perception because comprehension involves perception, acoustic-phonetic mapping, and lexical access (Koul, 2003). In fact, even for high quality synthetic speech, a substantial portion of cognitive resources are allocated to deciphering the acoustic-phonetic structure of the signal, leaving fewer resources available for higher level semantic processing (Duffy & Pisoni, 1992). Because a speeded sentence verification task is a relatively complex task, it is possible that participants' cognitive resources were focused on low level perception and thus unable to efficiently construct a mental representation of the message. If this were the case, such findings would have important implications for speech-generating-devices and for the individuals who use them.

4.2 Limitations

A limitation of this study is that true-or-false sentence pairs were used for the experiment stimuli. The 96 sentence pairs used were all altered SCOLP sentences. SCOLP format sentences were chosen because the SCOLP test has been proven to be a reliable measure of language comprehension (Adank & Janse, 2009), and because similar studies involving linearly time-compressed speech had used these sentences in the past (Adank & Janse, 2009; Adank & Devlin, 2010). Unfortunately, many of our participants reported that after they had gained some experience with the speeded sentence verification task, they realized that the sentence stimuli were arranged into pairs, (e.g. "Governors work in politics." vs. "Strawberries work in politics.") and that one member in the pair

would always be true and the other would always be false. This awareness enabled some participants to respond faster for the second sentence in a pair – they exhibited a repetition priming effect. The decrease in reaction time and increase in accuracy, which accompanied their repetition priming effect, meant that some participants performed better as they became increasingly familiar with the words used in the sentences, and with the sentences themselves. This effect counter-acted the decrease in comprehension that was predicted to occur as the speech signals became increasingly fast. If many words are initially recognized, then it is relatively easy to engage in a guessing strategy that reconstructs the initially unintelligible words (Dupoux & Green, 1997). Thus, for the second sentence in a pair it is possible that guessing strategies had a larger effect on response accuracy and reaction time than did actual signal comprehension. Future trials of this experiment could address this deficiency by continuing to use obviously true or false sentences for verification, but ensuring that each sentence occurs in isolation with no semantically related pair item.

Another possible limitation of this study is that the compression rates used were not as widely distributed as they perhaps should have been. Recall the five different linear time compression rates employed in this study: 42%, 40%, 38%, 36% and 34% of the original sentence durations. Dupoux and Green (1997) found that sentences compressed to 38% of their original duration were difficult for participants to understand, while Adank and Devlin (2010) found that listener's required 10–20 sentences to adapt to material that had been compressed to 35% of its original duration. In light of the contradictory past research, a pre-test was administered to 3 participants in order to determine a suitable range of compression rates. Participants in the pre-test heard eight sentences at each of the seven possible compression rates: 44%, 42%, 40%, 38%, 36%, 34% and 32%. Participants were seated in a quiet room and the sentences were played over a loudspeaker for all to hear. The pre-test participants exhibited excellent comprehension at the 44% compression rate and substantial difficulties in sentence comprehension starting at the 36% compression rate. On the basis of the pre-test participants' performance, it was decided that a 44% compression rate would be used for the training stimuli and that a 38% compression rate should be the median experimental compression value. We predicted that participant comprehension in Blocks 1 and 2 (42% and 40% compression) should be quite good as these two rates are slower than the median 38% value. We similarly predicted that comprehension in Blocks 4 and 5 (36% and 34% compression) should be quite poor as these two rates were faster than the selected median value. Surprisingly, the experimental participants in both speech conditions exhibited high comprehension throughout the experiment. Even in Block 5, the fastest compression rate presented, participants in the synthetic speech group still performed at above chance level (62.5%) for sentence verification accuracy. Future trials of this experiment could address this deficiency by using a broader

range of compression rates so as to better delineate the relationship between compression rate and comprehension. Furthermore, if a pre-test is administered, stimuli should be delivered in the same way (e.g. loudspeaker, headphones) as it will be delivered in the experiment.

5 Conclusion

In conclusion, our results have shown that despite recent advances in formant synthesis technology, listeners are still unable to comprehend synthetic speech as well as they comprehend natural human speech. Additionally, the comprehension of synthetic speech is more affected by adverse-listening conditions such as increased speech rate than is natural speech. Because text-to-speech generators play an important helpful role in the lives of visually and communicatively impaired individuals, and are widely used in the fields of language translation, business, and entertainment, these results are highly relevant. They indicate that further work is needed to improve the quality of synthetic speech for the sake of all individuals who use such signals. Our results thus add to the sizable body of research on synthetic speech perception. The researchers suggest that a similar study, which uses time-compressed speech to compare the quality of many different text-to-speech generators, would be beneficial to this field.

References

Adank, P. & Devlin, J.T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage* 49(1), 1124–1132.

Adank, P. & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *Journal of the Acoustical Society of America* 126(5), 2649–2659.

Baddeley, A.D., Emslie, H. & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing (SCOLP) test*. Bury St Edmunds: Thames Valley Test Company.

Boersma, P. & Weenink, D. (2009). Praat: doing phonetics by computer (Versions 4.5). [Computerprogram] http://www.praat.org/.

Duffy, S.A. & Pisoni, D.B. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language and Speech* 35(4), 351–389.

Dupoux, E. & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance* 23(3), 914–927.

Francis, A.L. & Nusbaum, H.C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attention, Perception, and Psychophysics* 71(6), 1360–1374.

Golomb, J.D., Peelle, J.E. & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *Journal of the Acoustical Society of America* 121(3), 1701–1708.

Janse, E. (2004). Word perception in fast speech: Artificially time-compressed vs. naturally produced fast speech. *Speech Communication* 42(2), 155–173.

Koul, R. (2003). Synthetic speech perception in individuals with and without disabilities. *Augmentative and Alternative Communication* 19(1), 49–58.

Koul, R. & Clapsaddle, K.C. (2006). Effects of repeated listening experiences on the perception of synthetic speech by individuals with mild-to-moderate intellectual disabilities. *Augmentative and Alternative Communication* 22(2), 112–122.

Koul, R. & Hanners, J. (1997). Word identification and sentence verification of two synthetic speech systems by individuals with intellectual disabilities. *Augmentative and Alternative Communication* 13(2), 99–107.

Koul, R. & Hester, K. (2006). Effects of repeated listening experiences on the recognition of synthetic speech by individuals with severe intellectual disabilities. *Journal of Speech, Language, and Hearing Research* 49, 47–57.

Mirenda, P. & Beukelman, D.R. (1987). A comparison of speech synthesis intelligibility with listeners from three age groups. *Augmentative and Alternative Communicaation* 3(3), 120–128.

Pallier, C. & Sebastian-Galles, N., Dupoux, E, Christophe, A. & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: A cross-linguistic study. *Memory & Cognition* 26(4), 844–851.

Schneider W, Eschman A & Zuccolotto A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Appendix

The number of syllables in each sentence is listed to the right of the sentence.

	Set 1			Set 2	
1.	Beavers build dams in the river.	8	1.	Governors work in politics.	8
2.	2. A tomato grows on a plant.	8	2.	Monks live in a monastery.	8
3.	3. Telephones can be bought in stores.	8	3.	Shovels are used in the garden.	8
4.	4. Motorcycles drive on the road.	8	4.	Sirloin steaks are sold by butchers.	8
5.	Fish breathe oxygen through gills.	7	5.	A leopard has a fur coat.	7
6.	Donkeys carry heavy loads.	7	6.	Butterflies have antennae.	7
7.	Carrots grow in a garden.	7	7.	A butcher works in a shop.	7
8.	An architect has a job.	7	8.	Wool is made from a sheep's coat.	7
9.	A camel is a kind of bird.	8	9.	Eagles build dams in the river.	8
10.	Dishwasher fluid walks the streets.	8	10.	A rainbow trout grows on a plant.	8
11.	Fathers are stored in the toolbox.	8	11.	Oxygen can be bought in stores.	8
12.	Biking is slower than walking.	8	12.	Fresh lemonade drives on the road.	8
13.	Buddhism is a pencil box.	8	13.	Pigs breathe oxygen through gills.	7
14.	Backpacks are always women.	7	14.	Babies carry heavy loads.	7
15.	Elephants are small insects.	7	15.	Beavers grow in a garden.	7
16.	April is a summer month.	7	16.	A vegetable has a job.	7

	Set 3			Set 4	
1.	A cake is baked in an oven.	8	1.	A tank is a weapon of war.	8
2.	Elephants are living beings.	8	2.	A minute is sixty seconds.	8
3.	Tables and chairs are furniture.	8	3.	Exercise is good for your health.	8
4.	Wooden chairs are for sitting on.	8	4.	A trout is a species of fish.	8
5.	Geese can fly long distances.	7	5.	A melon is a type of fruit.	8
6.	Bees fly around looking for food.	7	6.	Spoons are used for eating soup.	7
7.	A captain commands the ship.	7	7.	A shed is used for storage.	7
8.	Knives are used in the kitchen.	7	8.	Wine bottles are made of glass.	7
9.	A bike is a weapon of war.	8	9.	Strawberries work in politics.	8
10.	An hour is forty minutes.	8	10.	Donkeys live in a monastery.	8
11.	Smoking is very good for your health.	8	11.	A cake is used in the garden.	8
12.	An ant is a species of fish.	8	12.	Architects are sold by butchers.	8
13.	A cabbage is a type of fruit.	8	13.	A goldfish has a fur coat.	7
14.	Forks are used for eating soup.	7	14.	Bathroom sinks have antennae.	7
15.	Nurses are used for storage.	7	15.	A lion works in a shop.	7
16.	Policemen are made of glass.	7	16.	Ink is made from a sheep's coat.	7

Set 5

1.	A pelican is a bird species.	8
2.	Police officers walk the streets.	8
3.	Hammers are stored in the toolbox.	8
4.	Walking is slower than biking.	8
5.	Buddhism is a religion.	8
6.	Mothers are always women.	7
7.	An ant is a small insect.	7
8.	August is a summer month.	7
9.	Dentists are baked in the oven.	8
10.	Cabinets are living beings.	8
11.	The plastic doll is furniture.	8
12.	Computers are for sitting on.	8
13.	Grapes can fly long distances.	7
14.	Flies walk around looking for food.	7
15.	A leopard commands the ship.	7
16.	Snakes are used in the kitchen.	7

Do Japanese ESL learners' pronunciation errors come from inability to articulate or misconceptions about target sounds?

Akitsugu Nogita
MA Student, University of Victoria Linguistics
akitsugu@uvic.ca

This paper aims to examine whether Japanese English-as-a-Second-Language (ESL) learners' pronunciation errors are due to their inability to articulate, or to misunderstandings of target phonemes and the English phonological system. Four Japanese ESL learners read an English passage and some particular segments were analyzed for errors. After the analysis, each participant was interviewed about the errors. Results showed that the participants often purposely pronounced the same phoneme written with the same alphabet letter differently. For example, <v> in "gave" and that in "traveler" were misunderstood to be different phonemes. On the other hand, different phonemes spelled with different alphabet letters were often purposely pronounced the same. For example, <ar> in "hard" and <ir> in "first" were misunderstood to be the same phoneme. In addition, participants did not understand the whole picture of the English phonological inventory. I argue that the participants' mispronunciations are often due to the fact that they reportedly had not yet been taught basic English symbol–sound correspondence rules, not necessarily due to their inability to produce particular sounds. Since letter knowledge precedes phoneme awareness, the participants were not quite aware of English phonemes. If Japanese ESL learners in general adopt the same behaviour, pronunciation lessons need to pay more attention to Japanese ESL learners' understanding of the basic English phonological system, not only to what learners actually produce.

1 Introduction

This paper examines second language (L2) learners' understandings of the phonological system of a target language at the segmental level. The motivation of this study comes from the author's speculation that pronunciation errors by L2 learners are not necessarily errors, but instead may result from learners' own interpretations of L2 phonology due to learners' language backgrounds in which they have not yet been taught the phonological system of their target language. For example, if a Japanese English-as-a-Second-Language (ESL) learner

mispronounces "change" as [tʃɛndʒ][1] when it is supposed to be pronounced as [tʃendʒ], it would be careless to immediately conclude that the learner has difficulty in the distinction between the tense vowel [e], as in "pain," and the lax vowel [ɛ], as in "pen," because the learner's first language (L1) does not have this tense and lax distinction. There is a possibility that the learner has not yet been taught the basics of the English vowel system, and the learner misunderstands that "change" is supposed to be pronounced as [tʃɛndʒ]. If this is the case, the learner has yet to attempt to pronounce "change" as [tʃendʒ], and there is no way to know whether the learner has difficulty with [e]. Once the learner has been taught to pronounce "change" as [tʃendʒ], the learner might produce it without difficulty. Apparently, in the area of L2 pronunciation, analyses of learners' understandings of L2 phonology are less common than phonetic or phonological analyses of what learners actually pronounce. To address this gap in the research, this paper examines whether Japanese ESL learners have proper knowledge of English sounds when they commit pronunciation errors.

2 Background

2.1 Literature review

L2 learners' pronunciation errors are caused by factors other than difficulty in production. One possible factor is cognitive skills. For example, according to a finding of Fraser's (in press) study of /l/ and /ɹ/ distinction by Asian ESL learners, the participants produced these sounds without much difficulty, despite the fact that participants were unaware that /l/ and /ɹ/ were two different phonemes, which change lexical meaning in English.

Another possibility is what Richards (1971) called "false concepts" and what Stenson (1974) termed "induced errors" (as cited in Brown, 2007). These are errors caused by misleading teaching. In fact, Suarez and Tanaka's study (2001) with 88 Japanese college students found that 40% of the students claimed that their pronunciation problems came from a lack of pronunciation instruction in their six years of English curriculum in junior and senior high school. Another 24% felt psychological barriers had hampered correct pronunciation: when students try to pronounce English accurately, they are afraid of being teased or they feel embarrassed. From a teacher's perspective, according to Muroi's (2005) observation in Summer Teacher Training, about 30% of the Japanese teachers of English answered that they had never taught pronunciation to their students.

[1] Phonemic transcriptions of English vowels in this paper are based on "American English "R-Colored" Vowels as Complex Segments" Green (2001).

More specifically, Avery & Ehrlich (2003) explained an example of induced errors. They suggest that many mispronunciations by Portuguese speakers come from the influence of the Portuguese spelling system rather than from difficulty producing particular sounds; teachers with Portuguese students often familiarize themselves with the Portuguese spelling system. According to Bayraktaroğlu (2008), in terms of L2 learners' pronunciation errors, L1 orthographic interference and L1 phonological interference are completely different; the former is differences of one-to-one letter–sound correspondence between L1 and L2, while the latter is differences in the sound systems. The Japanese writing system has *Romaji*, Japanese Romanization, in which the symbol–sound correspondence rules are quite different from those in English in many respects. If Japanese learners of English familiarize themselves with the sound–spelling correspondences of Japanese Romanization similar to Portuguese learners of English, then their pronunciation errors may be a result of orthographic influence.

Moreover, L2 learners may need orthographic knowledge of the target language in order to understand its phonological system. Siegel and Wade-Woolley (1997) stated that phonological processing and literacy are strongly related. According to Carroll's (2004) study about first language (L1) acquisition, letter knowledge precedes phoneme awareness; letter learning helps children learn to separate phonemes from phonetic contexts and identify the same phoneme in different words. As well, according to Cook (2004), English speakers may understand that an alphabet character corresponds to an individual phoneme. When the number of alphabet characters and the number of phonemes do not match, for example "month" (five characters) and /mʌnθ/ (four phonemes), adult English speakers try to reconcile the contradictions in information, which results in difficulties in identifying the number of phonemes in the words (C. Pytlyk, in-class presentation). Furthermore, Goble (2002) revealed that his participants, Japanese college students, astonishingly lacked awareness that English loanwords in Japanese and their English counterparts are different entities, and the students' pronunciation and spelling errors showed an inordinate amount of loanword influence. Since Japanese has many loanwords from English, Japanese ESL learners' mispronunciations may come from loanword interference, rather than an inability to articulate. If Japanese ESL learners are not taught the English spelling system, they might be deficient in phonological awareness in English as well. It is worth examining this possibility.

2.2 Research questions

The present study is designed to address the following two questions:

1) When Japanese ESL learners make pronunciation errors, do they try to pronounce target sounds but fail to articulate the targets, or do they misunderstand or not know the target sounds to begin with? I specifically examine some segmental errors, namely the consonants /θ/, /ð/, /v/, /l/, /ɹ/, /f/, /h/, the distinctions between /s/ and /ʃ/, /t/ and /tʃ/, and /d/ and /dʒ/ before high front vowels, and the vowels /æ/, /ɛ/ and /e/.[2]

2) If Japanese ESL learners do not have knowledge of target sounds, what lessons do they need to understand target sounds properly?

3 Methods

3.1 Participants

There were four participants, labelled as P1, P2, P3 and P4. They were all Japanese ESL learners in British Columbia, Canada. Like the majority of Japanese people, all of the participants had studied English in junior and senior high school for six years. All of them claimed that they were not confident with their pronunciation, nor could they read the International Phonetic Alphabet (IPA), except P4, who was trying to learn the IPA on her own. P1 had been taught English by her mother who spoke British English, so P1's pronunciation might have been influenced by this exposure to the British accent. P1 was working, P2 was in a lower-intermediate class, and P3 and P4 were in an intermediate English class in an ESL school. Table 1 summarizes a number of the participants' traits.

Table 1. Summary of participant characteristics. Nb: Kansai= Osaka and Kyoto area. Kanto= the area around Tokyo.

	P1	P2	P3	P4
Age:	33	28	20	19
Gender:	F	F	F	F
Length of residence in Canada:	3 years	5 months	5 months	5 months
Home region in Japan:	Kansai	Kanto	Kanto	Kansai

[2] These segments are often considered problematic sounds for Japanese ESL learners (Avery et al, 2003; Ohata, 2004; Taniguchi, 2009).

3.2 Stimulus and procedure

The experiment took place over two days. On the first day, the participants read an English passage and Japanese nonsense words, and were asked to complete four phonemic contrast identification tasks. After their recordings were analyzed, I later talked to each participant individually about the results of the analysis. On the first day, the participants read the English passage, "The North Wind and the Sun" from the *Handbook of the International Phonetic Association* (p.44) (reproduced in the Appendix).

This reading task is designed to examine how participants interpret the segments mentioned in the research question 1. This story is familiar to many Japanese speakers and I expected that the participants would feel more comfortable with a familiar story than an unknown story. As well, this passage is commonly used in phonetic demonstrations. I handed the participants a sheet of paper with the passage on it a few minutes before recording, so they did not have time to ask native speakers about pronunciation or to check a dictionary. However, I taught them the sounds and meanings of presumably new words, such as "oblige" and "cloak." After practicing a couple of times, they were recorded. Recording was done in the Phonetics Lab at University of Victoria with a Luna 1.1 inch large diaphram condenser microphone, M-Audio Firewire 410, with PRAAT set to 44100 Hz.

Participants also recorded 10 nonsense Japanese words written with the Katakana syllabary. Table 2 shows the stimuli words presented to each participant. Some segments mentioned in Research Question 1, such as /l/ and /ɹ/, are obviously not distinctive phonemes in Japanese. Conversely, the contrasts in Table 2 are sometimes considered problematic (Avery & Ehrlich, 2003; Ohata, 2004) although these contrasts are also sometimes considered to exist in Japanese (Matsuzki, 1993; Inozuka, 2009). This task was designed to ascertain whether Japanese ESL learners have to articulatorily practice the distinctions between /s/ and /ʃ/, /t/ and /tʃ/, and /d/ and /dʒ/ before high front vowels, and the distinction between /ɛ/ and /e/, or if Japanese ESL learners can economically utilize L1 distinctions for these L2 distinctions.

Table 2. The 10 nonsense Japanese stimulus words. Nb: In Japanese [ti] and [tʃi] are distinctive, and so are [di] and [(d)ʒi] (Matsuzaki, 1993). Whether [ʃi] and [si] are distinctive is debatable; I argue that it depends on lexical classes.

ティー	[tiː]	チー	[tʃiː][3]		
ディー	[diː]	ジー	[(d)ʒiː]	ズィー	[(d)ziː][4]
シー	[ʃiː]	スィー	[siː]		
ベタ	[beta]	ベータ	[beːta]	ベイタ	[beita]

The words were aligned in this order on the sheet from which the participants read. They were asked to pronounce the words in a natural Japanese way. Since Japanese has phonemic pitch patterns,[5] which are not shown with regular orthography, most of the participants asked me about what pitch pattern they should use. Then, I answered that they could use whichever they felt was natural.

Participants also completed four phonemic contrast identification tasks. They were shown the homophones and minimal pairs in Table 3 and were asked to identify whether the words in each pair were the same or different in pronunciation. They were also asked to identify the difference between any two words they felt were not homophones. For example, I asked, "Do you know whether 'meat' and 'meet' are the same or different in pronunciation?" If participants confidently answered, "Yes, they are the same," I gave them a credit. If they showed uncertainty, "Um, I'm not sure. Maybe the same?" I did not give them a credit even if the answer was right as this may have been accidentally correct. The purpose of these tasks was to examine their L2 phonological and orthographic awareness in general. Since this task did not involve production, I could focus my investigation on the participants' understanding. The reason I chose these pairs is that each pair of words would likely be pronounced in the same way by Japanese speakers as a result of loanword adaptation processes. I tried to examine whether the participants could identify the phonemic structure of each word without being distracted by Japanese loanword adaptation.

[3] The transcription of the Japanese voiceless lamino-alveolo-palatal fricatives vary between [ʃ] and [ɕ] for the voiceless one, and between [ʒ] and [ʑ] for the voiced one (Pan, Utsugi and Yamazaki 2004). In this paper, I use [ʃ] and [ʒ] in order to be consistent with the English counterparts.

[4] In Japanese [dʒ] and [ʒ] are allophonic variations of one phoneme, and so are [dz] and [z] (Inozuka & Inozuka, 2009).

[5] Japanese is known as a pitch accent language in which pitch is the primary indicator of accent (stress) (Avery & Ehrlich, 2003).

Table 3. Phonemic contrast identification tasks

"meat"	versus	"meet"
"ear"	versus	"year"
"bone"	versus	"born"
"who'd"	versus	"food"

On the second day, I conducted semi-structured interviews with each participant in their L1, Japanese. Each interview lasted 30 to 40 minutes. I mainly asked participants what sound they actually tried to pronounce. This interview was designed to examine whether the participants' pronunciation errors were due to *mis-articulation of intended target* or *misunderstanding of target*. I used this method in order to best ascertain each participant's understanding in a straightforward manner. Below are examples of the questions:

- "Did you try to pronounce [θ] in 'north'?"
- "Did you try to pronounce [v] in 'gave,' 'of,' and 'traveler'?""
- "You pronounced [v] in 'gave' but pronounced [b] in 'of.' Can you come up with any reason?"
- "You said you did not try to distinguish /l/ and /ɹ/, and in fact you mostly did not. However, you pronounced native-like [ɹ] in 'wrap.' Can you come up with any reason?"

Below are examples of participants' answers:

- "Yes, I tried to pronounce [θ] to distinguish it from [s]."
- "No, I didn't put any extra effort into /θ/ and just pronounced [s] just as Japanese speakers commonly do."

When a participant mispronounced /θ/ as [s] and answered that she tried to pronounce [θ], her error was analyzed as mis-articulation of the intended target. When a participant mispronounced /θ/ as [s] and answered that she intended to pronounce [s], her error was analyzed as misunderstanding of target.

3.3 Analysis of the recordings

The transcribers were a phonetically trained native speaker of North American English and myself, a native Japanese speaker and trained phonetician. For the English data, the transcription was based on whether or not they produced the target phoneme; if a segment produced by a participant was obviously accented but was still easy to understand, it was considered correct. If a participant's production sounded like a different phoneme or was difficult to be categorized as any English phoneme, we transcribed what they actually pronounced. For the Japanese data, the transcribers judged what English phonemes could be represented by their productions.

4 Results and discussion

Overall, in terms of English phonemes that are generally considered problematic for Japanese ESL learners, participants quite often intended to pronounce different sounds. In other words, their mispronunciations often came from their understanding of what phonemes they were supposed to produce. More interestingly, the participants' own interpretations of English phonology varied much more than I expected from sound to sound, from word to word, and from individual to individual. Below are the details of each phoneme.

4.1 Results for /θ/

There are four occurrences of /θ/ in the passage, all of which are in the word "north." All the participants realized /θ/ as [s]. However, what is important in this paper is not the productions themselves, but whether the participants knew that the target was /θ/. For example, if P1 misunderstood that the target was /s/, instead of /θ/, she actually did not attempt to pronounce [θ]. In this case, I would conclude that she tried to pronounce [s] four times and successfully produced it four times. Consequently, what a teacher would then want to consider is teaching the proper target phoneme as it is identified orthographically, rather than articulation of [θ]. Therefore, I asked the participants if they knew that the target was /θ/, or if they intended to pronounce the dental fricative [θ].

- P1 reported she had never tried to pronounce [θ] although she knew that /θ/ and /s/ should be different. Therefore, she actually intended to pronounce [s] and she successfully produced what she was aiming to produce. In this case, there is no way to know if she was able to produce [θ] at the time of the recording because she had not attempted it.
- P2 reported she knew that the target was /θ/ and tried to pronounce it.
- P3 reported she knew that the letters <th> sounded different from the letter <s>, but she was more influenced by the English loanword in Japanese "ノース" [noːsɯ], which means "north." Since her underlying representation was /noːsɯ/, but not /nɔɹθ/, there is no way to examine if she was able to produce [θ] in this experiment.
- P4: Like P2, P4 reportedly tried to pronounce [θ].

The table below shows how many times the participants intended to produce the target phoneme and how many times they did so. Since P1 and P3 did not intend to pronounce [θ], the number of "Intended" is 0.

Table 4. Results for /θ/. Nb: "Occurrences: *T*"= The total occurrences in the passage; "Occurrences: Int"= How many times the participants intended to pronounce the target sound; "Correct: Int"= The number of correct productions when the participants intended to pronounce the target; "Correct: Acc"= The number of accidentally correct productions.

Target: /θ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	4	0	0	0	[s]: 4
P2	4	4	0	0	[s]: 4
P3	4	0	0	0	[s]: 4
P4	4	4	0	0	[s]: 4

4.2 Results for /ð/

There are 23 occurrences of /ð/ in the passage. The following summarizes participants' comments.

- P1: Just like [θ], she had never tried to pronounce [ð] and always realized it as the Japanese /z/ although she knew that /ð/ was not the Japanese /z/. In fact, she pronounced both [dz] and [z], which are allophonic variations of the Japanese /z/.
- P2 reported that targets were different from /z/ because the target was spelled as <th>, but not <z>. She said that she tried to distinguish between /ð/ and /z/ although she was not quite sure of the sound quality of /ð/. In fact, all of her /z/s were consistently pronounced as [z] or a somewhat devoiced [z]; she pronounced [ð] five times. She also pronounced the dental stop [d̪] nine times, which was close to the target /ð/ but was categorized as /d/. In fact, since the combination of phonological features [+continuant] and [−sonorant] is relatively difficult, stopping or affricating a fricative such as /ð/ to [d] is common in child language (Bernhardt & Stemberger, 1998).
- P3 reported definitely trying to distinguish between /ð/ and /z/. Moreover, she was aware that when function words, such as "the," "that" and "than" were followed by a difficult word in pronunciation, such as "traveler," she ignored the correct pronunciation of [ð] in order to concentrate her effort on the next word. Therefore, like P2, she understood the target correctly and she was aware of her errors.
- P4: Like P3, P4 reported trying to distinguish between /ð/ and /z/, which means she understood the target correctly. The big difference from P3 is

that she did not quite realize her production was not [ð], but its characteristic was much like the Japanese /z/.

Incidentally, all the participants did not know the cross-linguistic phonetic difference of /z/ between English and Japanese; the default form of J/z/ is the affricate [dz] (Grenon, 2008). Nevertheless, P1 produced the pure fricative [z] more often than [dz]. I will mention this phonetic issue in §4.13.

The table below shows how many times the target actually occurs in the passage, how many times the participants intended to produce the proper target, how many times the participants correctly pronounced the target when intending to do so, and how many times the participants accidentally pronounced the target. Also summarized are the incorrectly pronounced sounds of each participant. Since P4 missed the word "the" in the passage on one occasion, her occurrences were counted as 22. The question mark beside the number in the column of "Occurrences: Intended" means that the participant was not sure if she really intended to pronounce the target.

Table 5. Results for /ð/; Int= intended, Acc= accidental. Nb: In Japanese, [dz] and [z] are allophonic variations of the phoneme /z/ (J/z/) (Tsuzuki, 1996; Grenon, 2008; Vance, 2008; Inozuka and Inozuka, 2009). [d̪] is more dental than [d]. "?" indicates that the participants themselves were not really sure if they were aware of the target sounds.

Target: /ð/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	23	0	0	0	[z]: 20 [dz]: 3
P2	23	23?	5	0	[d̪]: 9 [dz]: 7 [z]: 2
P3	23	23	4	0	[d̪]: 8 [dz]: 8 [z]: 3
P4	22	22	0	0	[dz]: 11 [z]: 11 [ts]: 1

4.3 Results for /v/

There are six occurrences of /v/ in the passage: "gave," "of," and four occurences of "traveler."

- P1 reported she did not distinguish between /v/ and /b/, and substituted /v/ with the Japanese /b/. In fact, she pronounced both [β] and [b], which are allophonic variations of the Japanese /b/. Her /v/ in "of" sounded like [v], but it was actually the weakened version of /b/,[6] suggesting she accidentally hit the target.

[6] [β], [b] and [v] are all allophonic variations of /b/ in Japanese (Inozuka et al., 2009).

- P2: Due to having learned “gave” and “of” in Japan, P2 incorrectly memorized that “gave” and “of” were [geib(ɯ)] and [ob(ɯ)] respectively. Therefore, she misunderstood that the target sound was /b/, instead of /v/. As for “traveler,” she reportedly knew that there was [v], but she ignored the correct pronunciation of [v] in order to concentrate her effort on [l] and [ɹ].
- P3: As for “traveler” and “of”, she misunderstood that /v/ in these words was /b/. She explained that since she quite frequently used the word “traveler” in her conversation in Canada, she made up her own way to pronounce it. As for “of,” just like P2, she was not aware that “of” has a [v]. She pronounced [v] in one of the occasions of “traveler” but it was actually the weakened version of /b/, similar to P1. As for “gave,” she tried to pronounce [v] and successfully pronounced it. Therefore, she got one correct [v] out of one attempt.
- P4 reported possibly trying to pronounce [v] in “gave.” However, she actually pronounced this with a [b]. As for “traveler” and “of,” she consciously used the Japanese /b/. She mentioned that she concentrated too much on [l] and [ɹ] in “traveler” and could not afford to think about [v], just like P2.

Most of the time, the errors regarding /v/ were due to inappropriate intentions, rather than production problems. Only one instance, by P3, was appropriately intended and correctly produced.

Table 6. Results for /v/; Int= intended, Acc= accidental. Nb: “?” indicates that the participants were not really sure if they were aware of the target sounds.

Target: /v/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	6	0	0	1	[β]: 3 [b]: 2
P2	6	0	0	0	[β]: 5 [b]: 1
P3	6	1	1	1	[β]: 4
P4	6	1?	0	0	[β]: 5 [b]: 1

4.4 Results for /l/ and /ɹ/ in onset position

Japanese ESL learners often spend much time attempting to acquire the contrast between the North American English /l/ and /ɹ/ (E/l/ and E/ɹ/) because Japanese has only one liquid, /ɾ/, that can appear as [l] and even [ɹ] allophonically or in quasi-free variation (Magnuson, 2008). This distinction is so extensively studied that I put /l/ and /ɹ/ in the same section. From the point of view of this study, I

will give more credit to the participants who tried to distinguish between them but did not quite hit the target than to those who did not try to distinguish between them but accidentally hit the target. In examining whether the participants' attempts to distinguish /l/ and /ɹ/ affected the sound quality, I categorized their productions along two parameters following Magnuson (2008): *rhoticity versus laterality* and *central oral stricture*.[7] In this scheme, [l] is lateral and narrow while [ɹ] is rhotic and open.

There are 17 occurrences of /l/ in onset position, including consonant clusters, and 10 occurrences of /ɹ/ in onset position, including consonant clusters.

- P1 reported not trying to distinguish between /l/ and /ɹ/ in onset position at all. In fact, she almost consistently used rhotic liquid for both /l/ and /ɹ/. Interestingly, according to Magnuson (2008), J/ɾ/ is most commonly realized as a raised alveolar flap, but P1 pronounced [ɹ] much more often than a flap. The study by Akahane-Yamada, Aoyama, Fledge, Guion and Yamada (2004) showed that Japanese ESL learners more successfully acquire E/ɹ/ than E/l/ because the difference between E/ɹ/ and J/ɾ/ is perceptually more salient than the difference between E/l/ and J/ɾ/. In this way, P1 acquired E/ɹ/ and over-generalized it for /l/.
- P2 reported trying to distinguish between E/l/ and /ɹ/. In fact, she quite consistently pronounced more rhotic and open sounds, namely [ɹ] and [ɻ], for /ɹ/ while pronouncing lateral and narrow sounds, namely [l] and [ɭ], for /l/.
- P3 reported knowing that /l/ and /ɹ/ were supposed to be distinguished; however, she abandoned this contrast in her inter-language due to her low self confidence. In fact, she almost uniformly used rhotic and open sounds, namely [ɹ] and [ɻ], and the flap [ɾ] for both /l/ and /ɹ/. Interestingly, she pronounced an accurate [ɹ] in "wrap". However, she mentioned that the <w> in the spelling of "wrap" encouraged her to round her lips, which accidentally resulted in quite native-like [ɹ]. She was sure that she would pronounce the homophone "rap" with flap [ɾ]. Therefore, her [ɹ] was actually an accidental production caused by her misconception of the spelling and English phonotactics[8] where the sequence of /*wɹ/ at word-initial is not allowed.
- P4: Like P2, P4 said that she tried to distinguish between E/l/ and /ɹ/. In fact, she more successfully distinguished /l/ from /ɹ/ than P1 and P3. She pronounced [ɹ] better in "traveler" and "agree" than other words. She mentioned that she more frequently used the words "travel" and "agree"

[7] "Rhoticity" is "[ɹ]-like quality," while "laterality" is "[l]-like quality." "Central oral stricture" is how narrow or wide the space in the oral cavity is (Magnuson, 2008).

[8] Phonotactics deals with restrictions in a particular language on the permissible combinations of phonemes.

than “strong” or “around,” and she felt more comfortable pronouncing [ɹ] in familiar words.

Interestingly, P1 and P3, who did not try to distinguish /l/ from /ɹ/, performed notably worse with /l/ than P2 and P4, who tried to distinguish these. Conversely, P1 and P3 performed quite well with /ɹ/. Akahane-Yamada et al.’s finding that Japanese ESL learners acquire [ɹ] earlier than [l] might apply only to those who do not intend to distinguish /l/ from /ɹ/. Once they try, they might acquire [l] earlier than [ɹ] because [l] is less marked than [ɹ]. Moreover, English speaking children typically acquire [l] earlier than [ɹ] (Vihman, 1996).

Table 7. Results for /l/ in onset position; Int= intended, Acc= accidental. Nb: Ø= no consonant; [l] = the alveolar lateral; [ɺ] = the alveolar lateral flap; [ɾ] = the alveolar flap; [r] = the alveolar trill; [ɻ] = the lowered flap (the tongue does not quite touch the roof of the mouth); [ɹ] = the alveolar rhotic approximant.[9]

Target: /l/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	17	0	0	2	[ɹ]: 8 [ɾ]: 2 [ɻ]: 2 [ɺ]: 1 Ø: 2
P2	17	17	13	0	[ɺ]: 3 [ɻ]: 1
P3	17	0	0	1	[ɻ]: 6 [ɹ]: 5 [ɾ]: 5
P4	17	17	12	0	[ɾ]: 3 [r]: 2

Table 8. Results for /ɹ/ in onset position; Int= intended, Acc= accidental.

Target: /ɹ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	10	0?	0	9	[ɾ]: 1
P2	10	10	6	0	[ɾ]: 3 [ɻ]: 1
P3	10	0	0	4	[ɾ]: 4 [ɻ]: 2
P4	10	10	5	0	[ɾ]: 5

[9] According to the studies by Tsuzuki (1996), Magnuson (2008), and Inozuka et al. (2009), all of these sounds are possible allophones of J/ɾ/.

4.5 Results for velarized /l/ in coda position

- P1 reported not thinking about the difference between /l/ and /ɹ/ at all, so her correct production is considered to be accidental.
- P2 reported attempting to pronounce [l] but did not pronounce it successfully.
- P3 reported not attempting to distinguish between /l/ and /ɹ/. Just like onset position, she used the rhotic sound [ɹ] for /l/.
- P4 tried to pronounce [l] but actually produced [ɾ], which is neither rhotic nor lateral (Magnuson, 2008).

All in all, their realizations of /l/ in coda position seems the same as those in onset position; however, both P2 and P4 failed to pronounce [l].

Table 9. Results for velarized /l/ in coda position/; Int= intended, Acc= accidental, Ø= no consonant.

Target: /l/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	1	0	0	1	
P2	1	1	0	0	Ø: 1
P3	1	0	0	0	[ɹ]: 1
P4	1	1	0	0	[ɾ]: 1

4.6 Results for /ɹ/ in coda position

/ɹ/ occurs 20 times in coda position. I divided them into three smaller groups based on the preceding vowel: /ɔɹ/ as in "north," /ɑɹ/ as in "hard," and /ɝ/ in both a stressed syllable (as in "first") and an unstressed syllable (as in "stronger"). /ɔɹ/ occurs 8 times, /ɑɹ/ occurs once, and /ɝ/ occurs 11 times.

We will first examine /ɔɹ/ separately from /ɑɹ/ and /ɝ/ because the participants behaved interestingly. /ɔɹ/ occurs in three different morphemes, "north," "warm," and "more" in the passage, and all the participants consistently pronounced /ɔɹ/s in three different ways depending on the morpheme as shown in Table 10 below.

- P1 reported not really being aware that she pronounced "or" in "north" and <ar> in "warm" differently. However, she mentioned that she consciously pronounced <or> in "north" differently from the others because she was influenced by the pronunciation of Japanese teachers of English.

- P2 reported not knowing that <or> in "north," <ar> in "warm" and <or> in "more" were phonemically the same, and she was influenced by English loanwords. However, she was taught the pronunciation of "warm" by a native speaker in an ESL school, so she pronounced only "warm" and "warmly" correctly. Therefore, she was actually able to produce [ɔɹ] but misunderstood that <or> in "north" and <or> in "more" were not [ɔɹ]. In other words, she could not generalize the skill of pronouncing [ɔɹ] to words other than "warm."
- P3: As mentioned in §3.1, she pronounced "north" in the same way as the loanword [noːsɯ]. As for "warm", she misunderstood that <ar> in "warm" might be more like <ar> in "hard." As for "more," she pronounced it acceptably. However, she mentioned that, when facing <r>, she became intimidated and sometimes pronounced it strangely. Therefore, [ɔɹ] in "more" was counted as an accidentally correct production.
- P4 reported intending to pronounce the /ɔɹ/ in three different ways, consistent with what she had been taught in junior high school.

According to the participants' feedback, the three different realizations of /ɔɹ/ are not caused by phonetic environments. Rather, they are misconceptions that the /ɔɹ/ in all three instances was supposed to be different.

Table 10. Results for /ɔɹ/; Int= intended, Acc= accidental. Nb: "?" indicates that the participants themselves were not really sure if they were aware of the target sounds.

Target: /ɔɹ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	8	0	0	0	[oʊ] in "north" [oː] in "warm" [oa], [oə] in "more"
P2	8	2	2	0	[oː] in "north" [oa] in "more"
P3	8	0	0	2	[oː] in "north" [ʌː] in "warm"
P4	8	0	0	0	[oː] in "north" [aː] in "warm" [oə] in "more"

Note that the quality of the Japanese /a/ (J/a/) is between the cardinal vowels [a] and [ɑ], and it has a wider range of allophonic variations than the other Japanese vowels; [a], [ə], [ʌ] and [ɑ] can all be allophones of J/a/ (Tsuzuki, 1996). The participants produce J/a/ as [a], [ə], and [ʌ]. In addition, English loanwords in Japanese, "north," "warm" and "more" are typically adapted into [noːsɯ], [woːmɯ] and [moa] respectively.

Next, I will examine /ɑɹ/ and /ɝ/. Although /ɑɹ/ occurred only once in the

passage, it is interesting to compare /ɑɹ/ and /ɝ/.

- P1 reported not knowing that <ar> in "hard" and <ir> in "first" were different. In addition, she did not know that there was [ɹ] in those words (in a rhotic dialect). Although P1's mother, who has a British English background, taught her English when she was young, she did not know <ir> in "first" and <er> in "consider" or "other" were different in British English.
- P2: Like P1, P2 reported not knowing that <ar> in "hard" and <ir> in "first" were different. She said that she could produce the [ɹ]-like sound if she tried, as shown in §4.4 and §4.6, but she also insisted that [ɹ] in onset and coda sounded like completely different entities for her. This can be explained in terms of Brown's (2000) finding that Japanese subjects perceived E/l/ and E/ɹ/ in onset with only 31% accuracy while they did so in coda with nearly 100% accuracy (cited in Archibald, 2005). P2 also misconceived that the English letter <r> in coda position was the same as the Japanese symbol <ー>, which phonemically lengthens the preceding vowel. In fact, she quite consistently pronounced the long vowel [aː] for both /ɑɹ/ and /ɝ/. She misinterpreted the English orthographic information and did not know that English lacks the contrast of vowel length, unlike her L1.
- P3 reported not knowing <ar> in "hard" and <ir> in "first" were different. In fact, she pronounced [ɝ] in "hard" as she intended. However, as for the word "were," she said that she ignored [ɹ]. She pronounced "were" as [waː], as she intended. My interpretation of this is that she attempted /ɝ/ ten times out of 11 occurrences and succeeded four times. The problem is that she was not sure if /ɝ/ in each word in the passage was phonologically the same.
- P4: Like P2, she reported misconceiving that the letter <r> played the role of phonemically lengthening its preceding vowel. She also did not know that English lacks the contrast in vowel length present in her L1. In fact, she pronounced "hard" as [haːd] where [a] was lengthened, just as P1 and P2 did, while she pronounced <ir> "first" as [ɑɹ]. However, she believed that Japanese speakers can pronounce [ɹ], so that she was not intimidated by [ɹ]. She might have been aware that [ɹ] can appear as an allophone of J/ɾ/. The problem is that she did not know that /ɝ/ in each word in the passage was phonologically the same. In fact, the phonetic quality of her /ɝ/ varied from occasion to occasion even when her production was within the phonologically acceptable range. She mentioned that she pronounced /ɝ/ instinctually. Therefore, it may not be appropriate to consider that she attempted to pronounce /ɝ/.

All in all, the errors regarding post-vocalic /ɹ/ were mostly due to inappropriate intentions.

Table 11. Results for /ɑɹ/; Int= intended, Acc= accidental.

Target: /ɑɹ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	1	0	0	0	[aː]: 1
P2	1	0	0	0	[aː]: 1
P3	1	0	0	0	[ɝ]: 1
P4	1	0	0	0	[aː]: 1

Table 12. Results for /ɝ/; Int= intended, Acc= accidental.

Target: /ɝ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	11	0	0	0	[aː]: 9 [ɑː]: 1 [a]: 1
P2	11	0	0	0	[aː]: 9 [ə]: 2
P3	11	10?	4	0	[ə]: 4 [aː]: 1 [a]: 1 [oʊ]: 1
P4	11	0?	0	4	[aː]: 6 [ɑɹ]: 1

Again, [a], [ə] and [ɑ] are possible allophonic variants of J/a/. Note also that vowel length is phonemic in Japanese: e.g. /soɾi/ (sled) versus /soːɾi/ (Prime Minister).

4.7 Results for /f/ and /h/

It is important to note that Japanese has phonemic contrast between /f/ and /h/ which is neutralized before the vowel /ɯ/. Also, phonetically J/f/[10] is the bilabial fricative [ɸ] (Vance, 2008; Inozuka et al., 2009). In a questionnaire administered to 13 experienced ESL teachers in British Columbia, Canada, one respondent (and advanced ESL level instructor) pointed out the /f/ and /h/ distinction as one of Japanese learners' problems. As well, Berman, Lambacher, Martens, & Nelson (2001) found that Japanese learners perceptually confused /f/ and /h/ before [u]. Therefore, it is worth examining it. /f/ occurs five times and /h/ occurs eight times in the passage. The results show that the contrast between E/f/ and E/h/ does not seem problematic, except for "fold" and "who."

[10] Vance (2008) phonemicized the Japanese bilabial fricative as /f/. In this paper, I follow Vance's method. (cf. Akamatsu, 2000)

- P1 reported being aware of the cross-linguistic phonetic difference between E/f/ and J/f/. She pronounced the labio-dental [f], except <f> in "first" was [ɸ]. [ɸ] is more marked than [f] (Maddieson, 1984, 2005). Based on Eckman's Markedness Differential Hypothesis (2003), if one's L1 has a more marked sound, the less marked counterpart in L2 is not difficult to acquire. Therefore, her acquisition of [f] is not surprising although [ɸ] still appeared. She pronounced <f> in "fold" as [h], and she said that it was a slip of tongue. E/f/ is sometimes adapted as both J/f/ and J/h/ in loanwords: e.g. "telephone" can be pronounced and written as either /teɾefoN/ or /teɾehoN/[11] (Matsuzaki, 1992, 1993). She might have been influenced by that. As for "who," she did not know that <wh> in "who" and <f> in "food" were different. Therefore, she simply transferred L1 phonetics and phonotactics, namely neutralization of /h/ and /f/ before /ɯ/, and ended up with [ɸ] in "who."
- P2 reported not being aware of the phonetic difference between E/f/ and J/f/. Therefore, she was going to pronounce the bilabial [ɸ] and consistently did so. However, it was still within the acceptable range of E/f/. As for "who," she did not know that <wh> in "who" and <f> in "food" were different, like P1. J/f/ (or /h/) before a high back vowel was typically pronounced as [ɸ], similar to P1, but she happened to produce [h], or weakened [ɸ], in "who." Therefore, I consider it accidental.
- P3: Like P1, P3 reported being aware of the phonetic difference between E/f/ and J/f/. However, she pronounced "fold" as "hold," just as P1 did. The difference from P1 is that P3 more consistently produced [f] than P1, but she simply misread "fold" as "hold" and intended to pronounce "hold." As for <wh> in "who," she did not know that it was different from <f> in "food," like P1 and P2. Therefore, she simply transferred L1 phonetics and phonotactics, like P1.
- P4: Like P1 and P3, P4 reported being aware that E/f/ was not [ɸ], and she pronounced [f] in some words. However, she pronounced [ɸ] in "off." She said that she pronounced easy words, like "off," in the Japanese way, whereas she was careful with relatively difficult words. However, her [ɸ] was still phonologically within E/f/. The problem is that she purposely pronounced <f> in "off" and <f> in the other words differently, when English does not have this contrast. As for <wh> in "who," she did not know that it was different from <f> in "food," like all the other participants. P4 misconceived that <wh> in "who" was [f], and clearly pronounced "who" as [fu].

In some dialects, <wh> is categorized as /hw/ which is distinct from /w/. "Who" pronounced by P1 and P3 were phonologically within the acceptable range of

[11] /N/ stands for placeless moraic nasal.

such dialects. However, P1 and P3 did not intend to pronounce <wh> this way. In fact, they pronounced "when" as [wɛn]. Therefore, in this case, their L1 transfer happened to be within the acceptable range.

Table 13. Results for /f/; Int= intended, Acc= accidental. Nb: "?" for P4 is due to her purposely distinguished [f] and [ɸ] although English does not have this contrast.

Target: /f/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	5	5	4	0	[h]: 1
P2	5	5	5	0	
P3	5	4	4	0	[h]: 1
P4	5	5?	5	0	

Table 14. Results for /h/; Int= intended, Acc= accidental.

Target: /h/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	8	**7**	**7**	0	[ɸ]: 1
P2	8	**7**	**7**	1	
P3	8	**7**	**7**	0	[ɸ]: 1
P4	8	**7**	**7**	0	[f]: 1

4.8 Results for /t/, /d/ and /s/ before high front vowels

According to Ohata (2004), Japanese ESL learners may pronounce "seat" and "tip," for example, as like "sheet" and "chip" because they transfer the Japanese allophonic alternation of /t/, /d/, and /s/ which become [tʃ], [dʒ] and [ʃ] respectively before high front vowels. Such allophonic alternations occur in some classes of lexicon in Japanese; for example, the inflectional variations of the verb "win," /kata/ and /kato/(irrealis), /kati/(adverbial), /katɯ/(conclusive), and /kate/(imperative), in which the stem is /kat/, are pronounced as [kata], [kato], [katʃi], [katsɯ], and [kate] respectively. My focus is on whether this L1 transfer occurs at the level of their understanding or at the level of their production ability. I examined /s/, /t/ and /d/ before either /i/ or /ɪ/, namely "succeeded," "consider," "disputing," "did" and "immediately." Based on those allophonic

alternations, these words are expected to be pronounced as [səkʃidɛd] (or [səkʃidʒɪd]), [kənʃɪdɚ], [dʒɪspjutʃɪŋ], [dʒɪd] and [ɪmidʒiətli].

Contrary to the prediction, as shown in Tables 15 through 17, all the participants correctly pronounced these phonemes. In addition, participants said that the aforementioned predicted sounds were highly unlikely even in Japanese accented English, except [kənʃɪdɚ] for "consider" and possibly [səkʃid] for "succeed" were acceptable.

According to 15 scholars' interpretations of Japanese phonology in Matuzaki's (1993) paper, [ti] and [tʃi] are unanimously considered contrastive in Japanese except in some lexical classes mentioned above, and so are [di] and [dʒi]. Whether [si] and [ʃi] are contrastive in Japanese is debatable. Nogita (2010) argues that [si] and [ʃi] are not contrastive in core lexical classes, but they are in peripheral lexical classes such as technical terms and social dialects. As well, in Nogita's experiment, 93 monolingual standard Japanese speakers all distinguished [si] and [ʃi] regardless of their age. Thus, there is no reason that Japanese ESL learners have difficulty in pronouncing [t], [d], and [s] before high front vowels.

Additionally, the participants also recorded Japanese nonsense words written in Japanese orthography, and all of them distinguished "ティー" [tiː] from "チー" [tʃiː], "ディー" [diː] from "ジー" [(d)ʒiː], and "シー" [ʃiː] from "スィー" [siː] (as mentioned in §4.13). Therefore, if Japanese speakers pronounce /t/, /d/, and /s/, before high front vowels as [tʃ], [dʒ], and [ʃ], it makes more sense to consider that such errors are caused by other factors, such as loanword interference. In fact, there is variation in loanword adaptation. For example, [tɪ] in "tip" and [di] in "radio" were adapted as [tʃi] and [(d)ʒi] respectively, while "tea" [ti] and [dɪ] in "Disney" were adapted to [tiː] and [di] respectively.[12] The L1 transfer regarding [t], [d], and [s] is confined to loanword interference, but the transfer is not likely to occur in words that are not a part of Japanese vocabulary, such as "succeeded," "consider," "disputing," "did," and "immediately."

[12] Before the government in Japan standardized the writing system in 1991, the Agency for Cultural Affairs stipulated that in loanwords, [ti]/[tɪ] and [di]/[dɪ] in the original words should be written as "チ" [tʃi] and "ジ" [(d)ʒi] respectively as much as possible (with a few exceptions) (Matsuzaki, 1992).

Table 15. Results for /t/ before /i/ or /ɪ/; Int= intended, Acc= accidental.

Target: /t/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	1	1	1	0	
P2	1	1	1	0	
P3	1	1	1	0	
P4	1	1	1	0	

Table 16. Results for /d/ before /i/ or /ɪ/; Int= intended, Acc= accidental.

Target: /d/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	3	3	3	0	
P2	3	3	3	0	
P3	3	3	3	0	
P4	3	3	3	0	

Table 17. Results for /s/ before /i/ or /ɪ/; Int= intended, Acc= accidental.

Target: /s/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Accl	
P1	2	2	2	0	
P2	2	2	2	0	
P3	2	2	2	0	
P4	2	2	2	0	

4.9 Results for /æ/

As mentioned in §4.6, the Japanese /a/ is situated between the cardinal vowels [a] and [ɑ], and Japanese lacks a vowel in the low front region. I observed six occurrences of /æ/ in content words, namely “wrap,” “last,” and four occurrences of “traveller”. Since a vowel in a function word is often reduced to schwa, I did not include function words, such as “and” and “that.”

- P1 reported having never tried to pronounce [æ]. Despite her British background, she did not know <a> in "wrapped" and <a> in "last" were often pronounced differently in British English. However, she mentioned that she purposely pronounced "can't" as [kɑnt], instead of [kænt], even when talking to Canadian people because of her preference of British accent. At the same time, she realized that <a> in "can't" in British English was the same as J/a/, and in fact, the quality of her production had the characteristics of J/a/. Her production of /æ/ was a mixture of Japanese interference and over-generalization of British accent.
- P2 reported misconceiving that /æ/ in the passage was supposed to be the same as J/a/. Because of Japanese Romanization rule in which the letter <a> corresponds to the vowel J/a/, she had been habituated to this L1 symbol–sound correspondence.
- P3: Like P2, P3 also pronounced the letter <a> as J/a/ even in the English contexts.
- P4: She did the same as P2 and P3.

The results indicate that all of the participants did not try the low front [æ], but used J/a/.

Table 18. Results for /æ/; Int= intended, Acc= accidental. Nb: [a], [ʌ] and [ɑ] can be allophones of J/a/.

Target: /æ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	6	0	0	0	[ʌ]: 4 [a]: 2
P2	6	0	0	0	[a]: 3 [ʌ]: 2 [ɑ]: 1
P3	6	0	0	0	[ʌ]: 5 [a]: 1
P4	6	0	0	0	[a]: 5 [ʌ]: 1

4.10 Results for /e/ and /ɛ/

Ohata (2004) pointed out that Japanese ESL learners may make errors between the tense vowel /e/ and the lax vowel /ɛ/ because the Japanese vowel system does not have the tense-lax distinction. However, Ladefoged (2006) mentioned that the terms "tense" and "lax" are really just labels, as opposed to simply a matter of phonetic tenseness versus laxness. I will examine whether such errors come from Japanese ESL learners' misconceptions or their inability of production. /e/ occurs in "came," "they," "make," "take," and "gave," while /ɛ/ occurs in "when,"

"attempt," "confess," and two occurrences of "then." Since P1 misread "they" as "then," I counted four for the occurrences of /e/.

- P1 reported knowing that /e/ and /ɛ/ in the passage were different, but she also mentioned that she pronounced them "by instinct." She diphthongized /e/ and made it longer than /ɛ/. However, as long as she pronounced it "by instinct," the consistency may not be guaranteed.
- P2: She pronounced E/ɛ/ in the passage as the Japanese short monophthong /e/, likely because of the Japanese Romanization where it corresponds to the letter <e>. As for E/e/, she pronounced the target words as she was taught in junior high school. However, she was not sure that E/e/ in "came," "make," "take," and "gave" were the same as E/e/ in "they" because the spelling looked different.
- P3: She was sure that E/e/ in the passage was relatively diphthongized, while E/ɛ/ was relatively monophthongized. In fact, she distinguished them clearly in production.
- P4: She thought that E/e/ and E/ɛ/ in the passage were different, but it was because she had memorized those words with the Japanese pronunciation. She did not have connection to E/e/ and E/ɛ/ in the English phonological system. A lack of knowledge may be the reason of her mispronunciation.

Importantly, as I will mention in §4.13, the participants pronounced the nonsense Japanese words written in Japanese orthography: [beta], [beːta] and [beita].[13] The vowel part of the first syllable in each word was categorized as E/ɛ/, E/ɛ/ and E/e/ respectively, by the North American judge (see §3.3). The distinction of length did not change the English categories. Therefore, the fact that Japanese does not use the label of "tenseness" for grouping vowels does not mean Japanese speakers cannot distinguish between E/e/ and E/ɛ/. Since E/ɛ/ is one-mora and E/e/ is two-mora (Duran, 2005), Japanese speakers can pronounce these two by efficiently deploying the Japanese one-mora /e/ and the two-mora vowel sequence /ei/. However, E/e/ is usually adapted to three different Japanese categories depending on the words, namely J/eː/, J/ei/, and J/e/ – although J/e/ is not as common as the other two (Okada, 2004). For example, the English words "game," "paint," and "change" are adapted to /geːmɯ/, /peiNto/, and /tʃeNdʒi/. This inconsistent loanword adaptation may confuse Japanese learners of English. It makes more sense to consider that Japanese ESL learners' errors regarding E/e/ are because of loanword interference, not their inability to articulate.

[13] J/e/ is between the cardinal vowels [e] and [ɛ], so it can be transcribed as [ẹ].

Table 19. Results for /e/; Int= intended, Acc= accidental.

Target: /e/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	4	4?	4	0	
P2	5	5?	5	0	
P3	5	5	5	0	
P4	5	0?	0	0	[ɛ]: 5

Table 20. Results for /ɛ/; Int= intended, Acc= accidental.

Target: /ɛ/					
	Occurrences		Correct		Incorrectly pronunciations (sound: occurrences)
	T=	Int	Int	Acc	
P1	5	4?	4	0	[iː]: 1
P2	5	5?	5	0	
P3	5	5	5	0	
P4	5	5?	5	0	

4.11 Summary of segmental errors

Table 21 summarizes the numbers of errors due to production problems, errors due to inappropriate intentions, and accidentally correct productions. To see tendencies, I divided the errors into two types: "Consonants" (/θ/, /ð/, /v/, /l/, pre-vocalic /ɹ/, /f/, /h/, /s/, /ʃ/, /t/, /tʃ/, /d/, and /dʒ/) and "Vowels (Rhymes)" (/æ/, /ɛ/, /e/, /ɔɹ/, /ɑɹ/, and /ɝ/). As mentioned above, when the participants themselves were not sure whether or not they intended to pronounce appropriate targets, I marked this with a question mark in the corresponding results tables. In this summary I ignore those question marks.

Of the total 281 pronunciation errors, 186 (66.2%) were due to inappropriate intentions while 95 (33.8%) were due to production problems. Twenty-six productions that appeared to be correct were actually accidental. In detail, among 98 errors regarding vowels or rhymes, 92 (93.9%) were due to inappropriate intentions. This large number should not be ignored. In the errors regarding consonants, there are individual differences; P2 and P4 exhibited fewer inappropriate intentions than the others, and P2 and P4's total pronunciation errors were also fewer than the others'. Interestingly, in spite of P2's proficiency

in English being the lowest and that of P1 the highest, while P2 made the fewest errors in both. Based on these limited data, there seems to be no correlation between English proficiency and understanding of pronunciation.

Table 21. Summary of the results for the segmental errors: inappropriate intentions vs. production problems. Nb: Pro= the number of "production problems"; Int= the total number of "inappropriate intentions;" Acc= the number of "accidentally correct" productions.

	Consonants			Vowels (Rhymes)			Total		
	Errors		**Acc**	Errors		**Acc**	Errors		**Acc**
	Pro	**Int**		Pro	**Int**		Pro	**Int**	
P1	50		**13**	27		0	77		**13**
	1	**49**		0	**27**		1	**76**	
P2	37		**1**	24		0	61		**1**
	31	**6**		0	**24**		31	**30**	
P3	52		**6**	20		**2**	72		**8**
	19	**33**		6	**14**		25	**47**	
P4	44		**0**	27		**4**	71		**4**
	38	**6**		0	**27**		38	**33**	
T=	183		**20**	98		**6**	281		**26**
	89	**94**		6	**92**		95	**186**	

4.12 Results for the phonemic contrast identification tasks

According to the participants' comments and the summary of their errors, the participants seem to lack phonological awareness in English in many cases. In order to examine their phonological awareness more deeply, I asked the participants whether the homophone pair and minimal pairs were the same in pronunciation or not: meat/meet, ear/year, bone/born, and who'd/food. Since this task does not involve production, I could focus on the participants' understanding.

The result was that none of the participants were certain whether the words in each pair were the same or different in pronunciation. What is intriguing is that their production of both "meat" and "meet" sounded (almost) the same.[14] Nevertheless, the participants were still not certain that these words were

[14] Both "meat" and "meet" as loanwords in Japanese are also homophones: [mi:to].

homophones. Another interesting point is that between "ear" and "year," the pronunciation difference is obviously shown in the spelling, namely presence or absence of <y>, but none of them paid attention to it and became perplexed. The comments from each participant listed below are intriguing with respect to the participants' own interpretations of English phonology.

- P1: Between "who'd" and "food," she guessed that the tongue position might be different. (She did not mention for what sound the tongue position might be different.) What is interesting here is she paid attention to tongue position, rather than phonological categorization.
- P2 reported not knowing what the difference was, but she misunderstood that "meat" and "meet" were different because the spellings were different. Meanwhile, "bone" and "born" were the same because Japanese EFL learners typically pronounced these words in the same way, [boːN]. She inconsistently referred to either spelling or Japanese EFL learners' pronunciation or loanwords.
- P3 said that she had been pronouncing the two words in each pair probably in the same way, except she was taught that "ear" and "year" were different in junior high school although she was not sure what the difference was.
- P4 claimed that she had no awareness of the connection between spelling and sounds in English, or no knowledge about the English pronunciation system. In contrast, in Japanese she had the clear connection between orthography and sounds, and had the whole picture of the Japanese phonological system. Therefore, she had no idea about these English homophones and minimal pairs.

The participants' comments indicated that they do not really have a clear picture of the English sound system. Moreover, although they often referred to English loanwords in Japanese or the rules of Japanese Romanization, they did not fully depend on the Japanese phonology. Hocket (1960) defined linguistic sounds as discrete, whereas non-linguistic sounds form a continuum. More specifically, according to D. McKercher (personal communication, November, 2009), linguistic sounds must be categorized as phonemes in particular languages, while non-linguistic sounds cannot be categorized as phonemes. Since the participants often could not categorize sounds in the stimuli as particular English phonemes, they often might have pronounced English words with non-linguistic continuum sounds. In this way, the participants' vowel and rhotacized vowel qualities [15] varied substantially. It will be worth examining whether their vowel qualities will be more consistent after they learn the structure of English vowel inventory.

[15] "Rhotacized vowels" = /ɔɹ/, /ɑɹ/, and /ɝ/.

4.13 Orthographic pairs reading results

As mentioned earlier, when participants read Japanese stimuli written in Japanese orthography, they distinguished between [si] and [ʃi], [ti] and [ʧi], and [di] and [(d)ʒi]. As for their [e], [eː] and [ei], the first two can be categorised as E/ɛ/ while [ei] can be categorized as E/e/.

Interestingly, P1 (the participant with the longest residency in Canada at three years) showed different phonetic characteristics than P2, P3 and P4, whom had all lived Canada for five months in. P1 aspirated [t] in both a word-initial and word-medial position, and did not show pre-voicing for [d], [dʒ], or [b], and did not affricate [z]. In contrast, P2, P3, and P4 did not quite aspirate [t], (except P2 aspirated word-initial [t]) and often showed pre-voicing for [d], [dʒ], [b], and even [dz], and also pronounced the affricate [dz], instead of [z]. Japanese /z/ is typically the affricate [dz] (Grenon, 2008), as mentioned in 4.2 above. Vance (2008) mentioned that, according to some descriptions, Japanese /p/, /t/, and /k/ are typically weakly aspirated in word-initial position or in an accented syllable, and unaspirated elsewhere. According to Takada (2008), in Tokyo and Kansai region, voiced stops typically have negative voice onset time (VOT) values,[16] in other words "pre-voicing". Recall that P2 and P3 are from near Tokyo and P1 and P4 are from Kansai region. Therefore, P2, P3, and P4 showed typical phonetic characteristics in the Japanese stimuli, whereas P1 showed different characteristics. Since P1 had been in an English environment much longer than the others, her L1 may have been influenced by her L2. In fact, in Haraguchi's (2003) study, advanced Japanese ESL learners acquire English aspiration patterns without special endeavour. However, as shown in §4.1 to §4.10, P1's phonological realization was similar to that of the other participants. This implies that longer length of residence may help Japanese adult ESL learners acquire phonetic characteristics, but may not help them construct L2 phonological categorization. Incidentally, according to Hirayama (1994), Kyoto dialect speakers do not affricate /z/. Since P1 is from near Kyoto, her true fricative [z] is possibly not from L2 influence, but rather a characteristic of her L1 dialect.

Another interesting phonetic characteristic is that P1 and P4 added a glide in the Japanese [(d)ziː] and [siː] data, like [(d)zwiː] and [swiː], or unrounded [(d)zɯiː] and [sɯiː]. Conversely, they did not show such glide insertion in the English data. For example, they did not pronounce "succeeded" as [səksɯidɪd]. Their purposely differentiated productions may be due to orthographical interference. In the Japanese syllabary system, when a new syllable comes into use, it is written with the combination of two existing symbols (a big symbol and a small symbol), instead of creating a new symbol (Inozuka, 2009). For example,

[16] In the Tokyo area, younger speakers more often show positive VOT values in voiced stops than older speakers do (Takada, 2008).

the new syllables [(d)zi] and [si] are written with <ズィ> and <スィ> respectively. This two-symbol structure may cause some Japanese speakers to add an extra sound.

5 Discussion

5.1 Pedagogical implications

The pronunciation error patterns of these Japanese ESL learners can be divided into the following four types, summarized in Table 22.

Table 22. Summary of four types of the errors committed by the participants.

A:	A lack of phonological knowledge or misunderstanding of target sounds
B:	Abandonment of particular phonemes in learners' inter-language
C:	Difficulty in articulation or a lack of knowledge of the sound quality of a target phoneme
D:	Accidentally correct productions

Only C is a phonetic error, but the others are caused by misunderstanding. In fact, in many of the cases, the participants did not intend to pronounce the proper target phonemes. If native Japanese-speaking learners of English adopt the same behaviour, articulatory training often does not help them improve their pronunciation. The findings of this research suggest that pronunciation lessons need to stress learners' understanding of target sounds and the phonological system of the target language, and not only what learners actually produce. Each type of error is discussed in more detail below.

5.1.1 Type A: Lack of phonological knowledge/ misunderstanding target sounds

Learners do not know what they are supposed to pronounce. Learners often do not consider the target sound as a discrete phonological category, but as a non-linguistic sound. For example, the participants were not sure whether /e/'s in "gave" and "they" were the same. Therefore, their productions phonetically varied over a wide range. Another example is the participants' misunderstanding that /ɔɹ/s in "north," "warm," and "more" were supposed to be pronounced differently. The source of this type of error is that learners have not been taught the English symbol–sound correspondence rules. As Carroll (2004) stated, letter knowledge precedes phoneme awareness, as mentioned in 2.1. Learners need to know the concept of discrete phonological categorization with the visual cue, the

orthography. As well, as Makino (2008) mentioned, learners need to be shown all the English consonants and vowels to grasp the whole picture of the phonological inventory.

5.1.2 Type B: Abandonment of particular phonemes in learners' inter-language

Learners know what the target sound is, but they have abandoned the particular phoneme in their inter-language. For example, P3 knew that the English /l/ and /ɹ/ were different from the Japanese /ɾ/, but she gave up trying to acquire E/l/ and E/ɹ/, and substituted both with J /ɾ/. According to the participants' comments, they did not know why some particular phonemic contrasts, such as E/l/ and E/ɹ/, must be distinguished, and so they were not motivated to practice the contrasts. In order to help them understand the concept of contrasts, other ESL learners' errors or JSL (Japanese as a second language) learners' errors seemed effective. For example, naming the /p/ and /f/ confusion by Korean speakers and the /p/ and /b/ confusion by Arabic speakers, which are not problematic for Japanese speakers, helped the participants understand what the confusions like /l/ and /ɹ/ sound like.

5.1.3 Type C: Difficulty in articulation or a lack of knowledge of the sound quality of a target phoneme

Learners know what the target phoneme is and attempt to pronounce it, but fail to meet the target in terms of articulation, or can meet the target in isolation or careful speech but cannot afford it in a practical situation. Alternatively, they misunderstand the sound quality of the target phoneme. For example, P3 tried to pronounce /ð/ in the right place, but sometimes she affricated it. This type of error is a purely phonetic issue. Learners need some phonetic tips or need some practice on their own.

5.1.4 Type D: Accidentally correct productions

Learners accidentally met the target, but their production of target sounds was not intentional. In other words, it can be called a covert error. For example, P3 pronounced a target-like [ɹ] in "wrap", but actually she intended to pronounce /wɾ/, which happened to be realized as [ɹ]. Although this type of error may be difficult to find, the source of the error is the same as either A or B. Therefore, if learners understand the concept of phonological categorization and symbol–sound correspondence rules, this type of error can be avoided.

5.2 Limitations of the analysis

Firstly, this study's data were limited as some sounds appeared in only one word. For example, /θ/ appeared only in the word "North", and /ɑɪ/ appeared only once in the whole passage. There is no way to know how the participants pronounce these sounds in other words. However, this study was still able to identify the participants' understandings of these sounds, as mentioned in §4.1 and §4.6.

Secondly, the participants themselves were often not sure of what they intended to pronounce, which made it difficult to judge whether they attempted the proper target sounds or not. Moreover, they often used their L1 sound without understanding the L2 phonological inventory, and the L1 sound happened to be within the acceptable L2 target phoneme. For example, when they consistently used J/e/ for E/ɛ/, it sounded correct, but if they did not picture E/ɛ/, or so called "Short E", it is questionable whether they intended to pronounce the proper target or not. In these ambiguous cases, I put a question mark. At the same time, these ambiguous cases indicate that the participants did not understand the target sounds.

5.3 Future research

Firstly, this study found that the participants lack phonological and orthographical awareness in English in many cases. I should examine whether the finding from this qualitative study are generalizable by performing more quantitative research.

Secondly, P1 with three years of residence in Canada had better sound qualities of [ɹ] and aspiration than the other participants, while in terms of phonological and orthographical awareness, P1 was similar to the others. For example, P2 and P4 (with only five months of residence in Canada) performed much better than P1 in the /l/ and /ɹ/ distinction. Longer residence may help learners improve phonetic accuracy but may not help learners naturally acquire L2 phonological mapping and spelling rules. Since the participants had been exposed to English loanwords or strongly Japanese accented English for a long time, this exposure may have prevented them from constructing the L2 rules. This has to be studied with more participants in the future.

6 Conclusion

When Japanese ESL learners mispronounce English, they often intend to pronounce different sounds due to their misconception about target sounds, or due to their own interpretations of English phonology, as opposed to current ideas about Japanese learners' articulatory inability to produce particular sounds. Especially, in this experiment, the participants' errors regarding vowels or vowels followed by /ɹ/ were due to misguided intentions 93.9% of the time. Japanese ESL learners' misconceptions are likely due to their often not having been taught the basics of English phonological and orthographical systems. Therefore, the same phoneme spelled with the same alphabet letter(s) is often purposely pronounced differently when it appears in different words, For example, <v> in "gave" and that in "traveler" are misunderstood to be different phonemes. On the other hand, different phonemes spelled with different alphabet letters are often purposely pronounced the same: <ar> in "hard" and <ir> in "first" are in this way misunderstood to be the same phoneme(s). Moreover, even in the production of the notorious /l/ and /ɹ/, there was a clear difference in their productions between those who tried to distinguish them and those who did not. Insofar as Japanese ESL learners have not yet been taught the basic English symbol–sound correspondence rules, they would be limited to pronouncing according to their own interpretations of English phonology. If Japanese ESL learners have not yet pronounced their L2 sounds according to true English phonology, there is no way to know whether they actually have difficulty in producing particular sounds. It would be safe to avoid immediately concluding that pronunciation errors by Japanese ESL learners come from their articulatory inability to produce.

References

Agency for Cultural Affairs, Japan. (2002). *Kokugo-hyoki no kijun.* [Online resource]. Retrieved 30 Mar., 2009, from: http://www.bunka.go.jp/kokugo/.

Akahane-Yamada, R., Aoyama, K., Fledge, J.E., Guion, G. & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32, 233–255.

Akamatsu, T. (2000). *Japanese phonology*. Munich: Lincom Europa.

Akinaga, K. & Kindaichi, H. (Eds.). (2008). *Shinmêkai nihongo accent jiten*. Tokyo, Japan: Sansêdo.

Akiyama, K. (2009). *Sensê seminar: Hatsuon sassê*. NHK CTI Nihongo Centre.

Archibald, J. (2005). Second language phonology as redeployment of phonological knowledge. *Canadian Journal of Linguistics* 50, 285–314.

Avery, P. & Ehrlich, S. (2003). *Teaching American English Pronunciation*. New York: Oxford University Press.

Bayraktaroğlu, S. (2008). Orthographic interference and the teaching of British pronunciation to Turkish learners. *Journal of language and Linguistics Studies* 4(2), 107–143.

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.

Berman, J., Lambacher, S., Martens, W. & Nelson, B. (2001). Identification of English voiceless fricatives by Japanese listeners: The influence of vowel context on sensitivity and response bias. *Acoustical Science and Technology* 22(5), 334–343.

Bernhardt, B.H. & Stemberger, J.P. (1998). *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. California: Academic Press.

Brown, H.D. (2007). *Principles of language learning and teaching.* New York: Pearson Education.

Carroll, J.M. (2004). Letter knowledge precipitates phoneme segmentation, but not phoneme invariance. *Journal of Research in Reading* 27(3), 212–225.

Davenport, M. & Hannahs, S.J. (2005). *Introducing phonetics & phonology*. London: Hodder Education.

Duran, J. (2005). Tense/Lax, The Vowel System of English and Phonological Theory. In: P. Carr, J. Durand & C.J. Ewen (Eds.), *Headhood, elements, specification and contrastivity: Phonological papers in honor of John Anderson.* Philadelphia, PA: John Benjamins, 77–98.

Eckman, F.R. (2008). Typological markedness and second language phonology. In: J.G. Hansen Edwards & M.L. Zampini (Eds.), *Phonology and second language acquisition.* Philadelphia, PA: John Benjamins, 95–116.

Fraser, H. (in press). Pronunciation as categorization: The role of contrast in teaching English /r/ AND /l/. *Australian Review of Applied Linguistics*.

Gass, S.M. & Selinker L. (2009). *Second language acquisition*. New York: Routledge.

Goble, D. (2002). Loanword-induced interference in Japanese students' foreign language acquisition: Developing student awareness through experiential learning. *Journal of Chikushi Jyogakuen Junior College* 37, 55–77.

Green, A.D. (2001). American English "r-colored" vowels as complex segments. *Linguistics in Potsdam* 15, 70–78.

Grenon, I. (2008). *The acquisition of English soun[dz] by native Japanese speakers. A perceptual study*. Saarbrücken, Germany: VDM Verlag.

Haraguchi, Y. (2003). The acquisition of aspiration of voiceless stops and intonation patterns of English learners: Pilot study. *Proceeding of the 8th Conference of Pan-Pacific Association of Applied Linguistics*, 83–91.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics* 32(4), 565–589.

Hirayama, T. (1994). *Gendai nihongo hôgen daijiten*. Tokyo, Japan: Mêjishoin.

Hockett, C.F. (1960). The origin of speech. *Scientific American* 203, 88–96.

Inozuka, E., Inozuka, H. & Machida, K. (Eds.). (2009). *Nihongo onsêgaku no shikumi*. Tokyo, Japan: Kenkyûsha.

Ito, J. & Mester, A. (1995). Japanese phonology. In: J.A. Goldsmith (Ed.), *The handbook of phonological theory.* Cambridge, MA: Blackwell, 817–838.

Koizumi, T. (2003). *An introduction to phonetics*. Tokyo: Daigaku-Shorin.

Kurath, H. & McDavid, R.I. (1961). *The pronunciation of English in the Atlantic states.* Ann Arbor: University of Michigan Press.

Ladefoged, P. (2006). *A course in phonetics*. Boston: Thompson Wadsworth.

Maddieson, I. (1984). *Patterns of sounds*. New York: Cambridge Press.

Maddieson, I. (2005). Bilabial and labio-dental fricatives in Ewe. *UC Berkeley Phonology Lab Annual Report*, 199–215.

Magnuson, T.J. (2008). *What /r/ sounds like in Kansai Japanese: A phonetic investigation of liquid variation in unscripted discourse*. MA thesis, University of Victoria.

Makino, T. (2008). English pronunciation teaching and phonetic study. *Proceedings of the 162nd meeting of Osaka linguistics society.* Osaka, Japan.

Matsuzaki, H. (1992). A typification of variation in loanword sounds in dictionary items. *Tsukuba Working Papers in Linguistics* 10(11), 43–56.

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

Muroi, A. (2005). Hatsuon shido ni okeru kyôshi no yakuwari: Ayashii hatsuon shido no shôtai. *Êgokyôiku* 2005. 12. Taishûnshoten.

NHK Hosobunkakenkyûjo (Ed.). (2008). *Nihongo hatsuon accent jiten.* Tokyo, Japan: NHK Shuppan.

NHK Shuppan (Ed.). (2005). *NHK announce jissen training.* Tokyo, Japan: NHK Shuppan.

Ohata, K. (2004). Phonological difference between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL Journal* 6(4).

Okada, S. (2004). Language variation of the hiatus /ei/ in the Corpus of Spontaneous Japanese: In the case of loanwords. *IEICE Technical Report. Speech* 104(148), 35–40.

Ota, K., Riney, T.J., Takagi, N. & Uchida, Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35(3), 439–443.

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.

Siegel, L.S. & Wade-Woolley, L. (1997). The spelling performance of ESL and native speakers of English as a function of reading skill. *Reading and Writing* 9(5), 387–406.

Suarez, A. & Tanaka, Y. (2001). Japanese learners' attitudes towards English pronunciation. *Bulletin of Niigata Seiryo University* 1, 99–111.

Takada, M. (2008). Geographical pattern of VOT in Japanese initial voiced stops. *Studies in the Japanese Language* 4(4), 48–62.

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

The International Phonetic Association. (2003). *A guide to the use of the international phonetic alphabet*. Cambridge: Cambridge University Press.

Tsuzuki, M. (1996). An allophonic study of Japanese vowels and consonants. *The Journal of Aichi Gakuin University Humanities & Sciences* 43(2), 41–66.

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

Taniguchi, M. (2009). Japanese learners' weak points in English pronunciation, rhythm and intonation. Sênangakuin daigaku êgo shidoryoku kaihatsu wâkushoppu. [Online resource.] Retrieved 12 Aug., 2010, from http://www.seinan-gu.ac.jp/syakai/school/pdf/090822taniguti.pdf.

Vihman, M.M. (1996). *Phonological development: The origin of the language in the child*. Massachusetts: Blackwell Publishers Inc.

Yamaguchi, K. & Yamanaka, K. (2003). The confusion of alveolar and palato-alveolar fricatives by Japanese learners of English. *Bulletin of the Yamagata University. Cultural Science* 15(2), 207–220.

Appendix

Reading Task: The North Wind and the Sun

The north wind and the sun were disputing which was stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the north wind blew as hard as he could. But the more he blew the more closely did the traveler fold his cloak around him; and at last the north wind gave up the attempt. Then the sun shone out warmly, and immediately the traveler took off his cloak. And so the north wind was obliged to confess that the sun was the stronger of the two.

LOL! (laughing online): An investigation of non-verbal communication in computer mediated exchanges

Brittney O'Neill
University of Victoria
bkoneill@uvic.ca

Computer mediated communication (CMC) has a growing presence in modern communication. This paper discusses linguistic research in the field of CMC and explores how non-verbal communication, particularly instant messaging and texting, manifests in CMC. CMC has been found to be more conservative than speech but less conservative than written language. Emoticons and onomatopoeic expressions in CMC seem to play a role similar to non-verbal communication in face-to-face conversation; however, more research is necessary to confirm this.

Keywords: computer mediated communication, instant messaging, texting, non-verbal communication, emoticons

1 Introduction *(signing in)*

With the advent of personal computers and affordable internet connections, more and more people have begun to utilize computer mediated communication (CMC). These methods range from videoconferencing to forum boards to email, but this paper will focus on texting and instant messaging (IM). “Many teachers believe that students' wide use of "text speak" [is] a key factor in their students' negative performance” (Ross, 2007), claiming that this form of communication has negatively impacted the way students write. This paper will not focus on the concern that CMC is destroying the English language, but instead upon the potential problems of non-verbal expression in CMC. How can a speaker convey the depth of emotion present in non-verbal cues, such as facial expression and tone of voice, in a purely textual medium?

As a necessary precursor to further discussion, some important terms used in this paper must be defined. CMC shall be considered any method of communication between two or more sentient individuals occurring via interaction with a computational device (computer or cell phone) across a physical separation between interlocutors. For the most part, the form of CMC of

greatest interest will be IM, which constitutes a communication event that is text-based and synchronous, and which occurs between, typically, two participants, driven by a messaging client such as AIM, MSN Messenger, or Facebook Chat. Another important form of CMC is texting, which includes all textual messages sent via handheld devices to other handheld devices without any additional instant messaging client. Non-verbal communication shall be taken to encompass paralinguistic cues such as vocal tone and non-lexical sounds, proxemics, haptics, posture, eye contact, gestures, and facial expressions. In the context of CMC, traditional non-verbal communication is limited as a result of the spatial displacement of the participants; as such, for the purposes of this paper, the most relevant non-verbal communicators will be paralinguistic cues defined as onomatopoeic items and textual variation such as capital letters indicative of loudness, facial expression translated into emoticons or icons representing common facial expression, and gestures, also shown by emoticons.

2 Research and background *(finding contacts)*

Early forms of synchronous CMC arose in the 1960s (Huang, Yen & Zhang, 2008) but did not become widespread until 1993 in Europe with texting (Ling & Baron, 2007), and 1996 with the advent of ICQ in North America (Huang et al., 2008). Given the relative newness of CMC, it is not surprising that there is a fairly small pool of research. Not only is the available research fairly limited, it is also fairly diffuse. Researchers interested in the topic range from computer scientists to psychologists and sociologists, with a few linguists to round out the field. CMC research has begun to deal with a variety of common myths, including the idea that IM is primarily composed of abbreviations and that it is encouraging poor English. Less research has been done towards understanding how non-verbal communication translates into CMC. This paper will explore some of this research as well as the impact of the interaction between CMC and non-verbal communication.

Linguistic myths abound in many communicative domains; CMC is no exception. Some have even gone so far as to call CMC "the linguistic ruin of [the] generation" (Axtman, 2002, p. 1). Researchers have made numerous findings that discredit this notion. For example, Baron (2004) found that only 1.5% of words (in a large sampling of IM conversations) are replaced by stereotypical IM abbreviations (e.g. *l8r* 'later'), initialisms (e.g. *jk* 'just kidding'), and emoticons (e.g. ☺). Tagliamonte & Denis (2008) found a slightly higher rate of characteristic forms at 2.44%, but this was still a far lower percentage than expected. Given the popular conception that CMC is little more than acronyms and emoticons such as l8r <later>, kk <okay>, ttyl <talk to you later>, lol <laugh out loud>, brb <be right back>, and ☺, these findings are extremely significant.

If only 1.5-2.44% of words are being reduced to stylized IM speech, there remains a significant quantity of words used in more conservative, traditional linguistic expression. The reduction of *you* to *u* has also been observed to occur far less often than assumed by the general populace, with 91.4% of all instances of *you* retaining the full spelling and only 8.6% reducing to *u* (Tagliamonte & Denis, 2008). Ling & Baron (2007) also found similarly low percentages of shortenings and emoticon use in texting corpora. These results strongly suggest a need to reconsider the notion that CMC is reduced and "riddled with hieroglyphics, many of which [readers] simply could not translate"(BBC, 2003). It is, however, also important to note the sample size of the corpora used to draw these conclusions. Though comprised of many thousands of words, Tagliamonte & Denis's (2008) corpus (one of the largest used in this area of research) was comprised of the conversations of only seventy-one secondary school students. Until larger corpora are created, it is potentially deceptive to make generalizations. Seventy-one students are hardly representative of the entire community of IM users, particularly given that the secondary school students in question had volunteered to work with a mentorship program based out of the University of Toronto. These students are likely representative of a certain type of individual with interest in post-secondary education and academic research, and who may place greater emphasis on academic values than would other youth.

A number of researchers have also used their corpora to find similarities and differences between speech and CMC as well as between written language and CMC. In the same 2008 work, Tagliamonte & Denis concluded that, though not nearly as conservative as standard written communication, IM remains noticeably more conservative than speech in a variety of areas. That speakers choose to intensify only 12% of eligible adjectives in IM but 24% in speech suggests that IM tends towards a more conservative written-language style (lacking intensifiers such as *very* and *so*). However, given that there is a strong preference for more modern intensifiers (e.g. *really* and *so*) in IM, it must still be a more progressive speech-like form than standard written English (Tagliamonte & Denis, 2008). Similarly, Tagliamonte & Denis (2008) found language used in IM to be more conservative than speech but also more progressive than written English in the domains of quotative verbs (e.g. *be like* versus *says*), future references (e.g. *gonna* versus *will*), and deontic modals (e.g. *must, have to*). In all of these cases, IM seems to exist somewhere between the cutting edge of spoken communication and the fairly static standard written medium. CMC, despite being more conservative than traditional spoken communication, also seems to take a fairly pragmatic approach compared to many written standards. Texting and IM both tend towards using contracted forms, contracting 84.7% and 68.1% of potentials respectively (Ling & Baron, 2007), which mirrors a more speech-like or informal style. Texting also tends to omit apostrophes in contractions (likely due to complicated input methods), using only a third of required apostrophes (Ling &

Baron, 2007). Punctuation is also commonly dropped in both texting and IM with the sole exception of question marks, which appear following 73% of questions in texts and 100% of questions in IM (Ling & Baron, 2007). This sort of behaviour seems to support an efficiency and necessity approach to punctuation and may change as input technology improves. Unless there is the possibility of ambiguity without the punctuation – for example, a question being taken as a statement and thus not being answered – it is simply ignored for speed and ease of input.

Not only is CMC stylistically unique, it is also significantly removed from other communication mediums by physical manner restrictions. Where face-to-face conversation allows for a full range of non-verbal communication, phone conversations remove physical cues but still allow for paralinguistic cues, and signing conversations remove auditory/tonal cues but allow proxemic and visual cues. On the other end of the spectrum, letter writing lacks auditory and visual cues but tends to use far more advanced literary and rhetorical cues to evoke emotion. This occurs as a consequence of greater time to consider and revise utterances before sending. IM is like none of these. It lacks auditory and visual cues; and, being synchronous, requires relatively rapid responses, leaving less time for consideration and rephrasing. This is where onomatopoeic utterances and emoticons come in. Just as facial expressions and laughter punctuate face-to-face, phone, and sign language conversations at sentence and phrase breaks, utterances such as *haha* and *lol* appear in IM almost exclusively at these junctures, forming a similar sort of punctuation (Provine, Spencer & Mandell, 2007). This suggests that despite the significant differences in mode of expression, emotional cues are still present in IM. The implication is that emoticons and facial expressions may indeed be analogous forms, just as the textual onomatopoeias may be serving faithfully as soundless representations of paralinguistic cues. Nevertheless, a difference of intention does remain as IM emotional cues tend to be much more consciously expressed since they must be actively entered into a device, as opposed to face-to-face cues (Derks, Bos & Grumbkow, 2008). This allows for a more purposeful application of emotional quality than does an unconscious grin in face-to-face conversation. This, however, does not break the analogy suggested by the parallel patterning, allowing emoticons and onomatopoeia to be considered CMC analogues of non-verbal cues present in other types of speech. The presence of these iconic representations differentiates CMC from other written language forms, as most other forms use purely linguistic means to convey emotion without iconic emotional representation.

Given that emoticons and onomatopoeia seem to function as emotional cues in a purely textual environment, and given that emotional cues tend to be fairly common in face-to-face communication, it is striking that both Tagliamonte & Denis (2008) and Ling & Baron (2007) report such low percentages of emoticon

and onomatopoeia use in IM. This may be a result of the nature of the corpora or the possibility that volunteers chose not to submit more emotionally charged conversations. This again highlights the need for larger and more diversely constituted corpora of CMC to be developed in order to capture data that is representative of all manner of natural CMC, spanning age and conversation functions. Ethical issues may stand in the way of easy collection, but until a more substantial corpus is collected it will remain difficult to properly generalize trends, including those surrounding emotional expression.

Regardless of emoticons in IM, emoticon use has been found to correlate positively with enjoyment and degree of personal interaction perceived by users (Huang et al., 2008). It seems that, when "speakers" use emoticons, they feel more connected and also (perhaps as a result) experience greater enjoyment in the interaction. This may support the idea of emoticons as being analogous to facial expressions as, when people smile and behave expressively, they also tend to report greater enjoyment in interaction (Huang et al., 2008). On the other hand, people who are generally enjoying an interaction may use more positive emoticons, just as people who are enjoying a conversation tend to smile more. As such, it is hard to tease out the cause and effect relationship.

Derks, Fischer & Bos (2008) take a slightly different approach to investigating reasons for emoticon use by proposing, instead, that speakers use emoticons to clarify and intensify messages just as they use non-verbal cues in face-to-face contexts. In this view, emoticons play a pivotal role in communicative clarity and depth. Enjoyment is not primary in this theory. They also claim (in opposition to the findings of Tagliamonte & Denis, 2008) that "emoticons are used very often, especially in synchronous chat devices such as MSN" (Derks, Fischer & Bos, 2008). Though emoticons cannot elicit mimicry (Derks, Fischer & Bos, 2008), they do apparently fulfil the need for emotional contact, allowing CMC participants to develop intimacy and emotional connection to an equal or perhaps even greater degree than face-to-face speakers and listeners (Walther, 1995). In light of this research, it seems reasonable to suggest that emoticon use in IM may, in fact, be analogous to non-verbal communication in face-to-face contexts.

Beyond speakers' motives for using an emoticon or non-verbal display, it is important to consider the listeners' interpretation of the stimulus. Using artificial email messages (either positively or negatively inclined) with smiling, frowning, winking or no emoticons, Walther and D'Addario (2001) found emoticons unable to change the positivity or negativity of a message in the eyes of a reader. This does not support the idea of emoticons as equivalent to facial expression. These results may, however, be skewed by the strong valences of the statements which may have made them too absolute to be effected by the addition of a single non-verbal cue. Later work by Derks, Bos & Grumbkow (2008), with an added neutral condition and somewhat less absolute statements in both the positive and

negative conditions, found that emoticons were able to increase ambiguity and sarcasm when they contradicted the valence of the message. Smile emoticons also increased the positivity of a message that they were paired with, while negative emoticons increased negativity for positive and neutral conditions but not for negative conditions. This complicated interplay of emoticons and interpretation suggests that, like a physical expression of emotion, emoticons can help mediate message interpretation.

Both Derks, Bos & Grumbkow (2008) and Walther & D'Addario (2001) used artificial messages set in the context of an email (and not an IM). Though offering valuable insights, these research contexts somewhat reduce the generalizability of the results. Artificial messages are contrived, and as such may not accurately represent the cases wherein an actual speaker would apply emoticons to manipulate interpretation. The use of an email carrier also impacts the interpretations, as email is an asynchronous medium which allows for longer consideration and is less analogous to face-to-face conversation than is synchronous IM. Ideally, to fully investigate the non-verbal cues inherent in emoticons one must observe and test them in natural, synchronous conversational settings.

Though research has been done on a number of aspects of emoticon use and CMC, it remains incomplete and unable to fully explain the communicative application and impact of emoticons and onomatopoeia. The field remains open for future research.

3 Conclusion *(hitting 'send')*

CMC is becoming ever more prevalent in contemporary communication, and this increasing prevalence seems to be leading to decreased stigmatisation of the medium. Nonetheless, there are still those who believe that CMC, such as texting and IM, is an inadequate and even harmful method of communication. CMC is an area of interest that is, as of yet, fairly under-researched. It leaves plenty of room for new discoveries and insights and, as such, is incredibly inviting for young linguists. More and more people are building and maintaining relationships via CMC; subsequently, it is important to understand the differences between traditional letter writing, face-to-face communication, and CMC. How exactly we communicate emotion in a synchronous yet displaced and purely textual environment is a worthy field of study. Every time we gain new insight into the function of CMC we validate its use and prevent prescriptivist tendencies from hiding and even condemning the exciting new developments of the English language as it is used in a new medium of expression.

References *(contact list)*

Axtman, K. (2002). R u online?: The evolving lexicon of wired teens. *Christian Science Monitor*, 12 Dec., 2002, 1.

Baron, N.S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of Language and Social Psychology* 23, 397–423.

Derks, D., Bos, A.E.R. & Grumbkow, J.V. (2008). Emoticons and online message interpretation. *Social Science Computer Review* 26(3), 379–388.

Derks, D., Fischer, A.H. & Bos, A.E.R. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behaviour* 24, 766–785.

Huang, A.H., Yen, D.C. & Zhang, X. (2008). Exploring the potential effects of emoticons. *Information and Management* 45, 466–473.

Is txt mightier than the word? *BBC News*, 4 Mar., 2003. Retrieved from http://news.bbc.co.uk.

Ling, R. & Baron, N.S., (2007). Text messaging and IM: Linguistic comparison of American college data. *Journal of Language and Social Psychology* 26(3), 291–298.

Provine, R.R. (1993). Laughter punctuates speech: Linguistic, social and gender contexts of laughter. *Ethology* 95, 291–298.

Provine, R.R., Spencer, R.J. & Mandell, D.L. (2007). Emotional expression online: Emoticons punctuate website text messages. *Journal of Language and Social Psychology* 26(3), 299–307.

Ross, K. (2007). Teachers say text messages r ruining kids' riting skills. *American Teacher.* [Online resource, retrieved 14 Jun., 2010 from: http://findarticles.com/p/articles/mi_qa5369/is_200711/ai_n21298339/].

Tagliamonte, S.A. & Denis, D. (2008). Linguistic ruin? Lol!: Instant messaging and teen language. *American Speech* 83(1), 4–34.

Walther, J.B. (1995). Relational aspects of computer-mediated communication: Experimental observations over time. *Organizational Science* 6, 186–203.

Walther, J.B. & D'Addario, K.P. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review* 19(3), 324–347.

-s: The latest slang suffix, for reals

Vanessa McCumber
University of Victoria
vmccumber@alumni.uvic.ca

This paper explores the recent use of the suffix *-s* in English slang formation, as found in such forms as “whatevs,” “totes,” and “for reals.” A morphological analysis contrasts *-s* with similar suffixes and describes related morphological processes, including clipping and reduplication. The history, productivity and usage distribution of the suffix are also examined.

Keywords: *-s, slang, English, suffixation*

1 Introduction

The suffix *-s* serves many purposes in English: it can mark plurality, possession, or grammatical person. Recently it is also becoming popular for creating informal, slang versions of common words and phrases. For example, “see you later” becomes “lates!” and “this is certainly my favourite song” becomes “this is totes my faves song”. Including “for reals” in a sentence expresses the truth or sincerity of the statement.

This paper is exploratory and aims to describe the relatively new linguistic phenomenon of what will be referred to as “slang *-s*”, and will examine the suffix's morphology, history, and productivity, and its connections to other English morphological processes.

2 Data

Examples of words and phrases using slang *-s* are presented in Table 1. These were collected from the author's personal lexicon, self-reported data from colleagues and acquaintances (both male and female native English speakers, ranging in age from 12 to 60), and online sources. As slang, these words do not appear in any standard dictionaries, and, presumably because of their recency, only two were found in published slang dictionaries. Many words with slang *-s* can be found in Urban Dictionary, an online resource that is wholly user-edited and is thus much more current than most published works, but which lacks

professional standards and official citations. This website was used to find possible samples of slang *-s*, which were then confirmed in other sources before being included in the data set.

The data span several word classes, which suggests the productivity of slang *-s* is not limited to any one category. As discussed in the next section, though, there are likely two separate but related suffixes that constitute the slang *-s*.

The suffixation of *-s* does not appear to change word meaning, but rather serves to change the register to informal, colloquial speech; thus definitions are not included in the data set.

Table 1. Words and phrases with slang *-s* suffix.

Nouns	
babes < baby/babe	pops < pop
cutes < cutie	dins < dinner
homes < homey (< homeboy)	(what's) the haps < what's happening
Adverbs	
blates < blatantly	mabes/maybs < maybe
defs < definitely	obvs < obviously
for reals < for real	probs < probably
howevs < however	totes < totally
laters < later (< see you later)	supes < super
lates < later (< see you later)	whenevs < whenever
Adjectives	
brutes < brutal	faves < favourite
fabs < fabulous	perfs < perfect
Verbs	
jokes < I'm joking	
Interjections	
hells no/yeah < hell no/yeah	tevs < whatever
(oh) noes < (oh) no	wevs < whatever
no probs < no problem	whatevs < whatever
okays < okay	(oh) wows < wow
Pronouns	
whoevs < whoever	whomevs < whomever

3 Morphological analysis

3.1 Clipping

Slang *-s* suffixation is often, though not necessarily, accompanied by clipping, a process of "extracting a word from a longer word of the same meaning" (Millar, 2007, p. 39). When clipping occurs with slang *-s*, the word is often reduced to the first closed syllable, as in *totes* and *perfs. S*ometimes only the final syllable is dropped, as in *whatevs* and *howevs*, which may be because the monosyllabic form is blocked by another common word, e.g. *what-s* could be easily confused with *what's*, and thus it must be *whatevs* instead – this may also be why there are not many nouns that take slang *-s*, because the extremely common plural *-s* suffix would block the use of slang *-s*. In the case of words like *tevs*, the clipped form has been extracted from the middle of the word (*tevs* also demonstrates that there may be more than one slang form for a word, e.g. *whatevs*, *tevs*, and *wevs*). *Obvs* is an interesting example, since it ends in a consonant cluster, consisting of the coda of the first syllable and the onset of the following syllable. This may be because the *v* is perceived to be an integral part of the word, but may also be influenced by written English, where an abbreviated form of the word would likely maintain the *v* for clarity.

Slang *-s* could possibly be a form of embellished clipping (i.e. clipping plus suffixation), along the same lines as British *-ers* (e.g. *preggers* < *pregnant*, *starkers* < *stark naked*) (Huddleston & Pullum, 2002, p. 1636); however, the fact that some words with slang *-s* do not undergo clipping, such as *for reals* and *laters*, suggests that this is a stand-alone suffixation process that may optionally be accompanied by clipping, rather than a case of clipping accompanied by suffixation.

3.2 Hypocoristic

A review of the Oxford English Dictionary revealed that some of the words included in the slang *-s* data set are in fact derived from an older process of hypocoristic formation, dating back to the 19th century. The OED defines *-s* (suffix[2]) as a "shortened form of the hypocoristic diminutive suffix *-sy*, added to the same classes of words, as *babs*, *toots*; *ducks*, *moms*". This *-sy* suffix is a diminutive suffix added to proper names (e.g. *Betsy*, *Patsy*, *Nancy*) and common nouns (e.g. *ducksy*, *petsy*, *popsy*), and is also used with adjectives "expressing a degree of mocking contempt" (e.g. *artsy-fartsy*, *backwoodsy*, *folksy*) and may sometimes be considered a nursery form (e.g. *itsy-bitsy*, *teensy*) similar to words in *-y*. Huddleston & Pullum (2002) share this analysis, stating that *-s* may be used in terms of address (e.g. *pops*) or "playground words" with the diminutive *-ie*

(e.g. *onesies, widesies*). This playground suffix may be found on some of the same words as listed in the data set above; for example, *for realsies* and *favesies*.

In this way, the nouns included in Table 1, like *babes* and *homes*, may be separated from the other word classes and categorized under this older *-s* form rather than a new slang *-s* suffix. However, it is possible that the two suffixes interact with one another, and that the hypocoristic/diminutive form is the origination of slang *-s*, or that the rise in popularity of slang *-s* has renewed the use of the older *-s* form. More detailed analysis is necessary to determine exactly how the two suffixes are related.

3.3 Adverb *-s*

The OED includes another *-s* (suffix[1]) which is used in adverb formation and accounts for the alternations between null and *-s* in words like *anyway(s)* and *backward(s)*. It was originally a genitive suffix for Old English nouns and was used as an adverb-marker in Middle English. Although many slang *-s* words are adverbs, it is unlikely that the new suffix is derived from this older adverb-marking one, given their different functions (grammatical changing versus register changing) and the fact that slang *-s* can apply to many other word classes as well.

3.4 Additional morphological processes

Another word-formation process, *Mc*-reduplication, can be found to accompany slang *-s*. *Mc-* can be used as a prefix for forming (often pejorative) names, such as "Crazy McLegs", for a person with notable legs; alternatively "McGee" may be used to the same effect (e.g. "Legs McGee"). With slang *-s*, *McG-* is used as a prefix for rhyming reduplication to add an element of playfulness or emphasis, creating forms like *totes McGotes* (alternatively spelled *magotes*), featured in the film *I Love You, Man* (2009), and *brutes magutes*.

The suffix *-ski*, borrowed from Slavic languages, applies to the same classes of words as slang *-s*, leading to diminutive forms like *brewski* (< brew, = beer), and adverbs like *toteski* and *whatevski*. This suffix, like slang *-s*, does not have a specific meaning but rather changes the register to a more casual or playful one. In addition, *-ski* may then be followed by slang *-s* to create words like *toteskies* and *whatevskies*.

4 History and productivity

4.1 Time frame

Many words that take slang *-s* have only recently been used as slang. For example, in the Partridge Dictionary of Slang (2008), *fave* is only dated to 1921 (US usage), *for real* to 1952 (US), *later* to 1954 (US), *no prob* to 1971 (Australia), *perf* to 1979 (Australia), *totally* to 1982 (US), and *whatever* to 1989 (US). *Laters* and *lates* are listed with no date under the definition for *later*, but must be at least more recent than *later*'s date of 1954.

In Urban Dictionary, which has been online since 1999, most of the words with slang *-s* have been added between 2003 and 2005. This suggests that, although these words were almost certainly being used before these dates, they may have been increasing in frequency and becoming more accepted among a larger population in the early- to mid- 2000s.

A search of the Corpus of Contemporary American English (2008) revealed two print sources of slang *-s*: *whatevs* in a passage of dialogue in a young-adult fiction novel from 2006 (*Dial L for Loser*, Lisi Harrison), and *laters* in a short story in a 2008 issue of Asian-American magazine *Hyphen*. The British National Corpus (2001) has one undated recording of London teenagers using *laters* (13 conversations recorded by "Terry" (PS5A1)). There is also a recently published book titled *The Will to Whatevs* (Eugene Mirman, 2009). This scarcity of print sources confirms that slang *-s* is not yet well-established in mainstream media.

4.2 Usage distribution

According to self-reported data collected from native speakers, slang *-s* is available in at least some Canadian, American, and British dialects of English. It can be heard spoken in North American and British films and television, but is more readily found in online text. The Partridge Dictionary gives Australian origins to several clipped slang words, but not the forms in *-s*, so the suffix's presence in Australian English could not be confirmed.

Although present in spoken English, slang *-s* seems to be even more prevalent in written language, particularly in casual online writing and texting. The *-s* is occasionally spelled with <z> in some instances (e.g. *for realz* or *whatevz*), possibly in association with hip-hop culture where <z> is a common alternative to <s>, as seen in, for example, the film title *Boyz n the Hood* (1991).

4.3 Productivity

As seen in Table 1, slang *-s* can attach to a wide variety of word classes, including adjectives, pronouns, interjections, nouns (hypocoristic slang), and even verbs (e.g. the participle *joking* in *I'm joking*). The suffix seems least commonly attached to nouns, perhaps, as mentioned in Section 2, because of blocking from the plural *-s* suffix. It may be attached to a monosyllabic word, often in clipped form, or a bisyllabic one; the data are insufficiently complete to confirm the productivity for longer words. Except in the case of *noes* and *okays*, *-s* always attaches to a closed syllable. It is possible that these two words are part of a separate suffixation process, since they seem to be more like nursery or playground words, with the suffix *-ies*, rather than like the rest of the words in the data set.

Slang *-s*, when combined with clipping, could potentially be used for convenience, since the resultant words are shorter than the non-slang forms. However, there is sometimes a sacrifice in ease of articulation, for example in the word *obvs*, which has a cluster of three consonants with different places of articulation, which is rather difficult to say. This may be a case of writing influencing spoken word.

As a final note, many of the adjectives with slang *-s* are used as a verb complement rather than a noun modifier, as in “these cookies are fabs” rather than “these are fabs cookies”, although *faves* functions equally well as both.

5 Conclusion

This paper has explored the suffix *-s* as used to create informal slang words for colloquial speech. It is a recent development, dating no earlier than the mid- to late twentieth century, and gaining in popularity and use in the past decade. Slang *-s* is often accompanied by a clipping process, and is related or similar to other suffixation processes in English, such as diminutive *-sy* and playground *-ies*. It is productive for many word classes, including adverbs, adjectives and interjections, and is sometimes used in conjunction with *Mc*-reduplication and the suffix *-ski*.

Popular slang can quickly fall out of use, and given that slang *-s* has had a relatively short lifespan in which to establish a permanent place in the language, this suffix may eventually fade as a passing trend. However, slang words derived from clipping, to which slang *-s* is closely tied, appear to be fairly stable in the language, so perhaps slang *-s* too will become a standard component of the English word-formation repertoire.

References

BNC World. (2002). *The British National Corpus*, Version 2. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/.

Dalzell, T. & Victor, T. (Eds.). (2008). *The concise new Partridge dictionary of slang and unconventional English*. New York, NY: Routledge.

Davies, M. (2008). *The Corpus of Contemporary American English.* [Online resource, 10 Aug., 2010]. Available from http://www.americancorpus.org.

Hamburg, J.L. (Producer) & De Line, D. (Director). (2009). *I Love You, Man.* [Motion picture]. United States: DreamWorks Pictures.

Huddleston, R.D. & Pullum, G.K. (2002). *The Cambridge grammar of the English language*. Cambridge, UK: Cambridge University Press.

Millar, R.M. (2007). *Trask's Historical Linguistics* (2nd ed.). London, UK: Hodder Education.

-s, suffix[1]. (1989). In: *Oxford English Dictionary online*. [Online resource]. Retrieved from http://dictionary.oed.com/cgi/entry/50211174.

-s, suffix[2]. (1989). In: *Oxford English Dictionary online*. [Online resource]. Retrieved from http://dictionary.oed.com/cgi/entry/50211175.

-sy. (1989). In: *Oxford English Dictionary online*. [Online resource]. Retrieved from http://dictionary.oed.com/cgi/entry/50244843.

Urban Dictionary LLC. *Urban Dictionary*. (2010). [Online resource]. Available from http://www.urbandictionary.com.

Strategies for webpage trans-editing: A Socio-cultural study

Qian Xiaojuan
Ph.D. Student, University of Victoria Linguistics
xiaojuan@uvic.ca

Webpage trans-editing is a new trend in the study of translation variation. Based on previous research in news trans-editing and general translation studies, this paper firstly introduces some new types of trans-editing practice in order to redefine "trans-editing". From a socio-cultural perspective, a contrastive analysis of source texts and target texts is conducted to explore several strategies for webpage trans-editing. At the lexical and phrasal levels, this paper discusses the following strategies: domestication, foreignization, deletion, conversion, and explanation. At the discourse level, strategies of heading, selection, and restructuring are elaborated upon. In order to analyze the potential social and cultural factors involved in the choice of strategies in trans-editing, examples are selected from the internet, ranging from the areas of culture and politics to entertainment.

Keywords: trans-editing, webpage, translation variation, social and cultural factors, strategies, adaptation

1 Introduction

In Translation Variation Theory, Huang (2002) defines translation variation as an activity that aims at "deriving the relevant contents of the original version by using the appropriate adaptations[1]…in the light of specified readers' needs under specified conditions" (p. 35). Huang (2002) goes on to identify eleven translation varieties[2] including trans-editing (TE). Huang (2000) defines trans-editing as a special form of translation variation that involves the simultaneous interaction of both *translating* and *editing*. The source text (ST) is trans-edited such that the

[1] These include: expansion, deletion, editing, commentary, condensation, combination, and reformation (for more information refer to Huang, 2000, 2002).

[2] The eleven translation varieties: selective translation, translation and editing, translated narration, condensed translation, translated summarization, translated commentary, summarized translation commentary, explanatory translation, translation rearrangement, translation and writing, reference translation (for more information refer to Huang 2000, 2002).

most valuable information (to the particular readership) is reorganized in a sequence that is logical and acceptable to readers of the target language (TL). The emergence of new technologies has dramatically changed translation such that traditional "full" translation methods no longer meet the demands of the translation market. Insofar that TE is an efficient and time-saving type of translation since it selects only the information that is highly relevant to readers, it has become the focus of much academic attention in the field of Translation Studies (TS).

Mossop (2006) points out that the technological history of translation can be divided into pre- and post-computer eras. Webpage translation demands special translation techniques that are different from those involved in traditional, full translation. Trans-editing responds to this demand in that it is a very flexible and efficient translation strategy.

There have been several historical trends or approaches developed in TS that have viewed translation from different vantage points: linguistics (Catford, 1965), culture (Bassnett & Lefevere, 1990), ideology (Leung, 2006) and sociology (Wolf & Prune, 2007). In meeting the demands of localization within the larger context of globablization, trans-editing (and the trans-editing of webpages in particular) has involved more social and cultural adaptation.

This paper aims to analyze social and cultural factors in the choice of strategies for webpage trans-editing. Substantiated with examples from a number of official websites and the author's teaching material, this paper investigates the relevant strategies applied at the lexical, phrasal, and discoursal levels of translation. The findings are expected to inform real-life webpage trans-editing practice in English–Chinese and Chinese–English contexts.

2 Background

2.1 Trans-editing

Over the last decade, much of the research on trans-editing (Cheng, 2002; Huang 2000, 2002; Li, 2001; Sorby, 2008; Vuorinen, 1997; Xu, 1998) has focused on news trans-editing. Sorby (2008) conducts a comparative study of the complimentary versus derogatory word usage in the translation of news from English to Chinese. She points out that, while dealing with words that bear negative or positive connotation, translators usually transform the ST "through choosing the preferred meaning; changing behaviour, concept and mood; particularization; omitting or softening the critical tones" (p.25). In her article, "Bring the News Back Home: Strategies of Acculturation and Foreignization," Bassnett (2005) defines *news translation* as being different from the general term

translation, or traditional full translation. She also claims that news translation displays a stronger tendency toward acculturation.[3]

Many scholars in TS from China have contributed to the research on news trans-editing from English to Chinese or the other way around. Among them, Liu and Huang (2001), and Xu (1998) have addressed the basic principles involved in TE; Chen (2009) has discussed the ideology-related norms in English–Chinese news trans-editing, and Zhang (2004) has analyzed the trans-editing process from a functional point of view. To the extent that much TE research has focused mainly on news and journalism, this paper will first attempt to address the trans-editing of other text types besides news.

2.2 Webpage trans-editing

With the advent of the Internet, translation has become an activity that is closely linked to computer networks. Scholars of information technology tend to tackle technological problems in their research on translation, while those from the translation field itself focus on the theoretical and strategic aspects of network translation. Tang and Gentzler (2008) point out that translation has become a network-supported commercial process because of globalization. At the same time, webpage translation, such as that of government portal websites or the websites of multinational corporations, needs to be localized in order to provide target readers translations that they find highly acceptable.

Webpage translation is a diverse enterprise, which covers an array of topics including (among others) economics, politics, science, technology, culture, and entertainment. Consequently, the translation of webpages requires knowledge of various fields and translation strategies to cope with different text types. However, due to the constraints of time and space as well as regulatory constraints, webpage translation has unique features. For example, unlike full translation, less relevant information (relative to the readership) can be omitted. The layout of a webpage can also be redesigned to cater to the target culture, and information can be reshuffled for the convenience of target readers.

[3] Acculturation is the eradication of traces of "otherness" in a text so as to reshape the text for domestic consumption in accordance with the norms and expectations that prevail in the target culture. It brings a text more completely into the target culture since that text is effectively aimed at readers with no knowledge of any other culture (Bassnett, 2005).

2.3 The definition of trans-editing

Huang (2000) groups trans-editing (originally named "translation and editing") together with another ten translation varieties (see footnote 2) as a major type of translation variation. TE has gone far beyond the field of news. Based on a review of the literature, I have come across the following types of trans-editing:

i. webpage trans-editing (see section 2.1.2),
ii. classical works trans-editing, such as the adapted versions of Shakespearean works or Bible stories for children,
iii. summary reports, such as synopses of academic papers in TL for general readers,
iv. chart trans-editing, such as transcribing graphical business information into statements in the TL, and
v. movie reviews trans-edited from reviews in the SL, which may even include the translator's own opinion of the movie.

In this way, trans-editing can be redefined as a variety of translation, which involves the deployment of various adaptation skills to select and restructure information, to transform text types, and to edit culturally sensitive issues in order to cater to the needs of various readerships. Trans-editors, who have been likened to tailors, act as editors or filters to pick up or trim off information in order to adapt a social or cultural reality in the ST into a corresponding reality in the target text (TT). Compared with fully translated texts, trans-edited works are free in style and form, and satisfy the particular needs of a target readership.

3 Methodology

This paper investigates strategies applied in webpage trans-editing and aims to identify socio-cultural factors that may impose constraints on the choice of trans-editing strategies. All the English data are selected and retrieved from the following websites: The *New York Times*, *British Broadcasting Corporation (BBC)*, *The Daily Telegraph*, and *eBeijing*. The Chinese versions are from *People's Daily Online*, *Sohu*, or from students' revised work. Most of the examples are displayed first in the source language, then in the target language. The Chinese versions have been augmented with glosses. For the convenience of contrastive analysis, different translation versions and back translation versions (hereafter: BT) are also provided. The analysis mainly focuses on socio-cultural factors, such as cultural traditions, historic events, political attitudes, and logical sequencing.

As Venuti (1995) says that translation processes are mediated by the cultural values of the TL in a hierarchical order, the analysis here is conducted first at the

lexical and phrasal levels, and then at the discourse level. Note that these levels interact with each other within the larger external social-cultural frame.

Since trans-editing at the discourse level involves different versions of source texts, all the texts are attached in the appendices. Due to the focus of the analysis and the length of the target text, back translation is not provided for the full Chinese text but for the title and subtitles, which are vital to restructuring.

4 TE Strategies at the three levels

Strategies for webpage trans-editing have undergone many adaptations from the traditional translation strategies since the advent of online networks. What are the strategies frequently taken by webpage trans-editors? Chen (2009) points out that cultural particularity, target readership, and ideology may have effects on the choice of trans-editing strategies. What is the influence of socio-cultural factors on the choice of strategies in webpage trans-editing? How are the strategies carried out in the practice of TE? Previous research has addressed strategies but has done so without systematic categorization or practical exemplification. For example, strategies at the discourse level have been theoretically proposed by Huang (2000, 2002) and Xu (2003), but without detailed illustration. The following discussion is conducted at the lexical, phrasal and discourse levels to exemplify what the different strategies are and how those strategies are applied in webpage trans-editing.

4.1 Lexical and phrasal levels

At the lexical and phrasal levels, translators (and especially trans-editors) need to be aware of connotative and symbolic meanings since these are the components of pragmatic meaning (Sophia, 2000). Based on the research of Huang (2000, 2002) and my own previous research, I will discuss the following trans-editing strategies frequently adopted at the lexical and phrasal levels: domestication, foreignization, deletion, conversion, and explanation.

4.1.1 Domestication vs. foreignization

Venuti (2008) has discussed two translation strategies: foreignization and domestication. The former, which is more salient in translationese,[4] retains the "foreign" features of the ST while the goal of the latter is to reproduce the ST author's intended meaning in an idiomatic and natural style appropriate to the TL.

(1)

七夕
(*qi xi*)

Gloss: seven night

Translation 1: Chinese Valentine's Day

Translation 2: Double Seventh Day

In the culture section of the official Beijing Government website there is an article introducing the traditional Chinese festival "七夕 (*qi xi*)," which is held on the seventh day of the seventh month of the Chinese lunar calendar. "Qi xi" is translated as "Chinese Valentine's Day," which is a typical example of domestication (from the perspective of the target readership). Since the purpose of the website is to attract tourists and provide some basic cultural information, the partial mapping between the traditional Chinese festival and the Valentine's Day of the West will help the target readers respond to the text in a similar way as Chinese readers do. In that way, the otherwise unfamiliar Chinese festival will be brought closer to the target audience because it has been acculturated to suit its readership.

Conversely, "Double Seventh Day" is an example of the use of *foreignization* to attract the target readers' attention to the unfamiliarity of Chinese culture, and thereby pique their interest. In this way, this strategy can prompt readers to further explore relevant Chinese customs, as well as idioms and sayings developed out of the festival.

[4] Translationese: Translations exhibit their own specific lexico-grammatical and syntactic characteristics. These "fingerprints" left by the translation process were first described by Gellerstam and named translationese (Alexander, 2010).

4.1.2 Explanation

To bridge the cultural gaps between the SL and the TL, trans-editors can provide explanations by adding further details within brackets or footnotes, or by transplanting cultural aspects into the translation itself. However, they must be cognizant of the acceptability of the translation and the patience of target readers: the explanation needs to be precise and concise, and avoid providing detailed background information.

(2)

中秋节
(*zhong qiu jie*)

Gloss: mid autumn festival

Translation: *Mid-Autumn Festival: A Time for Reunion*

In the same article from the Beijing Government website, the term 中秋节(*zhong qiu jie*) is translated as "Mid-Autumn Festival: A Time for Reunion." The translation has adopted the strategy of explanation. The phrase following the colon is additional information compared with the source term. For those who are unfamiliar with Chinese culture, the word "mid-autumn" conveys nothing but the concept of time. However, the added "a time for reunion" pinpoints the significant feature of this traditional Chinese festival. The added information can help target readers bridge the cultural gap between the SL and TL, and induce a similar response to the target readership. In this way, the explanation assists the target readers' understanding of the source information.

4.1.3 Deletion

Information loss is inevitable in translation because of cultural gaps. A target text may lack certain culturally relevant features that are present in the source text; therefore, loss in translation is "due to the fact that the backgrounds, shared knowledge, cultural assumptions and learnt responses of monolingual TL speakers are inevitably culture-bound" (Rastall, 1994, p.40). Deletion is adopted when no culturally corresponding terms exist or because of sensitive political issues. As Seidman (2006) points out:

> "…[T]ranslation is necessarily also a political negotiation, it appears not strictly as a linguistic exercise but also in a variety of relational modes: translation as colonialist, imperialist, or missionary appropriation but also translation as risk, as assimilation, as treason, as dislocation, as survival." (p.9).

Translation has long been perceived as a political act. To work within the constraints of publishing rules and governmental regulations, and to avoid sensitive issues or to soothe targeted readers, trans-editors usually take political factors into consideration. Trans-editors are different from traditional translators because they enjoy more freedom even though they have to trans-edit according to the requirements or needs of the readership.

(3)

> At that time, China was in the throes of Chairman Mao Zedong's bloody Cultural Revolution.

Even though the Cultural Revolution (1966–1976) has brought serious nation-wide disaster and turmoil in modern Chinese history, Chinese people are still grateful to and respectful of Chairman Mao for his extraordinary contributions. Therefore, the original word choice "bloody," which is derogative, will not be emotionally accepted in the Chinese context. Searching Google, there is no such combination "Mao Zedong," "bloody" and "the Cultural Revolution" in Chinese. Therefore, when trans-editing the original sentence, trans-editors may delete "bloody" from their texts to soften the tone and to avoid irritation among the target readership.

4.1.4 Conversion

With respect to complimentary and derogative language in Chinese, Sorby (2008, p.19) notices "single English words which may have a positive or negative connotation, depending on their context, are translated into Chinese using separate terms." For example, the simple English word "die" can be translated into Chinese complimentarily as "牺牲(*xi sheng*—sacrifice one's life; die a martyr's death; lay down one's life)、献身(*xian shen* – give one's life to; devote one's life to)、仙逝(*xian shi* – pass away)," or derogatively as "毙命(*bi ming* – violent death)、暴毙(*bao bi* – sudden death)、蹬腿儿(*deng tuier* – kick the bucket)." Sorby goes on to point out:

> "[C]omplimentary and derogatory terms play such a large role in the Chinese language, it is very important for translators to make the correct judgment when using the strongly connoted words in their translation in order to avoid either creating an unexpected effect or leading to serious misunderstandings" (p. 23).

In TE, trans-editors may change the original complimentary/derogative feature of the lexical term in order to cater to the ideology of target readers.

(4)

China 'patriot' sabotages auction

Translation: 佳士得巴黎拍卖圆明园兽首 神秘中国买家竞拍成功
(*jiashide bali paimai yuanmingyuan shoushou, shenmi zhongguo maijia jingpai chenggong*)

Gloss: Christie Paris auction the Imperial Summer Palace animal heads mysterious Chinese buyer auction success

BT: Christie auctioned the Chinese Sculptures in Paris. A mysterious Chinese successfully got the bid.

Example (4) is the translation of the title of a report from the BBC's website. The story concerns the auction of two Chinese sculptures held by Christie's Auction House in Paris, which was won by a Chinese bidder. "Sabotage" is a rather sensitive word. According to online Oxford English Dictionary, "sabotage" means "the malicious damaging or destruction of an employer's property by workmen during a strike, any disabling damage deliberately inflicted, especially which carried out clandestinely in order to disrupt the economic or military resources of an enemy." It is "malicious" and "clandestine." To achieve a better title for the TT, the trans-editor needs to be clear about the story of this auction, as it is stated in this report: "[The sculptures] were originally taken by British and French troops from the Imperial Summer Palace in October 1860 towards the end of the Second Opium War. China had tried to stop the sale, and later threatened the business of Christie's in China for having gone ahead." To most Chinese given this context, the behaviour of the Chinese bidder is patriotic since it stops the illegitimate auction and enables the return of the looted sculptures. Hence, "sabotage" is not translated derogatively as "破坏 (*po huai* – destroy)" but as "竞拍成功 (*jingpai chenggong* – auction success)," which means a successful bidding.

4.2 The discourse level

4.2.1 Restructuring

One of the shortcomings of webpage information is that some articles are not carefully edited due to time constraints. To reproduce the source message in the target text, trans-editors need to restructure the text and omit irrelevant or

redundant information in order to meet the particular needs of target readers. For example, Bible stories or simplified versions of Shakespearean works are published for children. Irrespective of whether the task involves intra-textual or inter-textual trans-editing[5], trans-editors need to restructure or reorganize the TT such that the TT is consistent with the ideology of the TL. The Appendices A and B together provide a good example of inter-textual trans-editing.

Appendix A consists of three articles in memory of Gerald Ford from *The Daily Telegraph* and *New York Times*. Appendix B is the TT trans-edited from Appendix A, with reference to students' translation practice along with a trans-edited article from the website www.people.com.cn. The STs have been systematically restructured into a Chinese text (Appendix B(i): "The Legendary Life...") with four clearly named sub-sections (B(ii–v)). Various aspects of Ford's life have been selected from the three articles and grouped into the four different sections which are headed by distinct titles: (ii) the Unprecedented, (iii) the Embarrassing, (iv) the Lucky, and (v) the Controversial aspects of his life story. The following examples illustrate the treatment of each of the four section titles in Appendix B.

B. ii.

史无前例：未经选举当上副总统和总统
(shi wu qianli wei jing xuanju dangshang fu zongtong he zongtong)

Gloss: history no previous example: no election be vice president and president

BT: Unprecedented: President and Vice-president without Election

B. iii.

尴尬往事：摔下"空军一号"
(*ganga wangshi shuaixia kongjun yi hao*)

Gloss: embarrassing past events: falling off Air Force One

BT: Embarrassing: Falling Down from Air Force One

[5] Intra-textual and inter-textual trans-editing are two major trans-editing types classified by Huang (2002). The former refers to trans-editing of a single text while the latter refers to multi-text trans-editing.

B. iv.

福星高照：躲过两次暗杀
(*fuxing gao zhao duoguo liangci ansha*)

Gloss: lucky star shine high: escape over two times assassinations

BT: Lucky: Surviving Two Assassinations

B. v.

成也“水门”，败也“水门”：赦免尼克松丢总统宝座
(*cheng ye shuimen bai ye shuimen shemian nikesong diu zongtong baozuo*)

Gloss: success be Watergate, failure be Watergate: pardon Nixon lose president throne

BT: Controversial: Watergate – Pardoning Nixon and Losing the Presidency

4.2.2 Headings

Headings, or titles, which may constitute the front page of a website, are similar in function to journalistic headlines. A news headline, if effective, must meet the requirements of capturing the essence of the event and attracting readers' attention.

To illustrate this strategy, the following examples from different websites were chosen from various subject areas like economics, health, and entertainment.

(5)

Mum Doesn't Live with Us Any More
(More than 150,000 UK mothers live apart from their children as courts increasingly give custody to fathers, Catherine Bruton reports on the rise of 'mothers apart.'...)

Translation: 离婚率攀升，“超级奶爸”涌现
(*lihun lv pansheng chaoji naiba yongxian*)

Gloss: divorce rate increase, super daddy spring up

BT: Increase in Divorce and Popularity of “Super Nanny Daddy”

The source headline focuses on the absence of mothers and the lead (i.e., the most important structural element of a story) elaborates on the phenomenon that more children are living with their fathers. The headline in the TT combines the information from the headline and the lead from the ST, and supplies the term "super nanny daddy." As a title on a webpage, "super nanny daddy" (coined from "super nanny") is very eye-catching to the target Chinese readership.

(6)

Anne Hathaway to play Judy Garland on film, stage

Translation: "公主" 踏上 "绿野仙踪"
(*gongzhu tashang lvye xianzong*)

Gloss: "princess" steps on "green field and celestial track"

BT: "Princess" on her way to "Wizard of Oz"

The actress, Anne Hathaway, is remembered by the audience for her performance in the movie *The Princess Diaries*. Judy Garland is the actress who starred as Dorothy in *The Wizard of Oz*. Ms. Hathaway is going to portray Ms. Garland in both film and on stage. The creative adaptation in the headline has greatly increased the acceptability of the TT because the masterpieces of the two actresses are more provocative than their exotic names to the target Chinese readers.

4.2.3 Selection

To restructure the ST, trans-editors need to know how to select the most important information from it.

(7)

福特总统的传奇人生
(*fute zongtong de chuanqi rensheng*)

Gloss: Ford president *-de* legendary life

BT: The Legendary Life of President Ford

The title in the TT "福特总统的传奇人生 (*fute zongtong de chuanqi rensheng*)" suggests that the TT's perspective is that of Ford having lived an eventful life. The four subtitles in the TT are guideposts for trans-editors' selection of only those points that are closely relevant to what each subtitle suggests.

Section one, which is centered on Ford's political life, covers information about his Presidency, Vice-Presidency, and overall political contribution to the U.S.A. Section two is about his embarrassing experiences and negative views of him due to the military withdrawal from Vietnam and dismal economy. The third part is about his personality and his luck. The last section is about the key issue of Ford's life: his controversial act of pardoning Nixon, which directly resulted in his loss of the Presidency. By selecting and reorganizing the major relevant information, trans-editors can produce a TT that is appropriately organized for its audience.

This paper will not discuss other aspects of webpage trans-editing, such as webpage design, layout, or animation since these aspects are not the concern of translators but of webpage designers and administrators.

5 Conclusion

Webpages reflect much of the social and cultural aspects of and among societies. To ensure the efficient and successful transmission of information, translators have adopted trans-editing in response to market demands for time-effectiveness and localization of target texts. At each of the lexical, phrasal, and discourse levels, the choice of trans-editing strategies is influenced by socio-cultural factors such as the particular needs of the readership including its conventions, political and cultural norms, and ideology.

At the lexical and phrasal levels, similar to full translation, information can be added or deleted, and the complimentary or derogatory characteristics of a word or phrase may be changed in accordance with the ideology of the target readership. At the discourse level, trans-editors may use strategies of restructuring, adding headings, or selection to efficiently and successfully convey the source information. Hopefully the strategies that have been discussed and illustrated here can inform the work of trans-editing practitioners and student translators.

References

Bassnett, S. & Lefevere, A. (Eds.). (1990). *Translation, history and culture*. London: Printer.

Bassnett, S. (2005). Bringing the news back home: Strategies of acculturation and foreignisation. *Language and Intercultural Communication* 5, 120–130.

British Broadcasting Corporation. (2009). China 'patriot' sabotages auction. BBC News website, 2 Mar., 2009. [Online resource]. Retrieved 9 Apr., 2009 from http://news.bbc.co.uk.

Chen, Y. (2009). Quotation as a key to the investigation of ideology-related norms in English-Chinese news trans-editing. *Compilation and Translation Review* 2, 1–36.

Cheng, M. (2002). The principles and strategies of trans-editing for the news media. *Journal of Translation Studies* 7, 113–134.

Chinaculture.org. (n.d.). Chinese Valentine's Day. [Online resource at: eBeijing, the Official Website of Beijing Government]. Retrieved 29 Nov., 2009, from http://www.ebeijing.gov.cn/Culture/EnjoyBJ/t950627.htm.

Folaron, D. (2007). Introduction: Connecting translation and network studies. *Meta* 52(4), 605–642.

Gerald Ford dies; Nixon's successor in '74 crisis was 93. *The New York Times*, 27 Dec., 2006. [Online resource]. Retrieved December 28, 2006, from http://www.nytimes.com.

Harnden, T. (2006). The man who healed wounds of Watergate. *Telegraph*, 28 Dec., 2006. [Online resource]. Retrieved 28 Dec., 2006, from http://www.telegraph.co.uk.

Huang, Z. (2000). *On translation variation (in Chinese)*. Beijing: China Translation and Publishing Corporation.

Huang, Z. (2002). *Translation Variation Theory (in Chinese)*. Beijing: China Translation and Publishing Corporation.

Key moments in the life of Gerald Ford. *Telegraph*, 28 Dec., 2006. [Online resource]. Retrieved 28 Dec., 2006, from http://www.telegraph.co.uk.

Li, D. (2001). Adaptive translation of international news from English to Chinese: An exploration of methods. *Journal of Translation Studies* 6, 47–60.

Liu, L. & Huang, Z. (2001). The basic principles for trans-editing: Strategies for translation variation. *Chinese Science and Technology Translators Journal* 14, 42–44.

Mossop, B. (2006). Has computerization changed translation? *Meta* 51, 787–805.

Nida, E.A. (2001). *Language, culture and translating*. Shanghai: Shanghai Foreign Language Education Press.

Rastall, P. (1994). Communication strategies and translation. *Babel* 1, 38–48.

Seidman, N. (2006). *Faithful renderings: Jewish–Christian difference and the politics of translation.* Chicago: The University of Chicago Press.

Sophia S.A.M. (2000). *Pragmatic meaning and cognition*. Netherlands: John Benjamins Publishing Company.

Sorby, S. (2008). Translating news from English to Chinese: Complimentary and derogatory language usage. *Babel* 54, 19–35.

Tang, J. & Gentzler, E. (2008). Globalization, networks and translation: A Chinese perspective. *Perspectives: Studies in Translatology* 16, 169–182.

Venuti, L. (2008). *The translator's invisibility: A history of translation.* London/New York: Routledge.

Vuorinen, E. (1997). News translation as gatekeeping. In: M. Snell-Hornby, Z. Jettmarova & K. Kaindl (Eds.), *Translation as intercultural communication: Selected papers from the EST Congress, Prague 1995*. Amsterdam: John Benjamins, 161–171.

Wines, M. (2009). China's leaders see a calendar full of trouble. *The New York Times*, 9 Mar., 2009. [Online resource]. Retrieved 10 Mar., 2009 from http://www.nytimes.com/2009/03/10/world/asia/10china.html?_r=1.

Wolf, M. & Prune, E. (2007). Introduction: The emergence of a sociology of translation. In: M. Wolf & A. Fukari (Eds.), *Constructing a sociology of translation*. Amsterdam: John Benjamins, 1–36.

Xu, J. (2003). Book review: Huang, Z. (2002): Translation variation theory (in Chinese), Beijing, China Translation & Publishing Corporation. *Meta* 48(4), 590.

Xu, M. (1998). An overview of trans-editing. *Chinese Translators Journal* 6, 30–34.

Xu, M. (2003). *English news and translation*. Beijing: China Translation and Publishing Corporation.

Zhang, M. (2004). Theory and practice of adaptation: A case analysis within the functional theory of translation. *Journal of Sichuan International Studies University* 20, 95–100.

Appendix A

i. The man who healed wounds of Watergate

(Harnden, 2006)

Gerald Ford, who died on Boxing Day aged 93, was hailed by President George W Bush yesterday as a "man of complete integrity" who helped to bring America together after the Watergate scandal.

Mr Ford was unexpectedly thrust into the United States presidency and worked to heal what he called the "poisonous wounds" left by Richard Nixon's resignation. He had done so "with common sense and kind instincts", Mr Bush said.

In his 895-day presidency he pardoned Mr Nixon – an act that probably cost him the presidency in 1976 when he was narrowly defeated by Jimmy Carter – and ended US involvement in Vietnam.

"On August 9, 1974, he stepped into the presidency without ever having sought the office," Mr Bush said. "He assumed power in a period of great division and turmoil. For a nation that needed healing and for an office that needed a calm and steady hand, Gerald Ford came along when we needed him most."

Mr Bush said his fellow Republican "reflected the best in America's character" and helped restore confidence in the integrity of the presidency through the "honourable conduct" of his administration. One of his unexpected legacies was the rise to prominence of several politicians who came to shape Mr Bush's presidency.

Both Vice-President Dick Cheney and Donald Rumsfeld, who recently resigned as defence secretary, served as Mr Ford's chief of staff.

The only White House occupant never elected either to the presidency or the vice-presidency, Mr Ford became the longest living former president. A former congressman, his highest ambition, he said, had been to become Speaker of the House of Representatives.

His greatest contribution during his presidency, he believed, was to help unify his country after the Nixon era. At his inauguration he declared that "our long national nightmare is over" and vowed to "bound up the internal wounds of Watergate, more painful and more poisonous than those of foreign wars".

His most controversial act was his first one – the pardoning of Mr Nixon "for any crimes he may have committed". There were accusations this was part of an arrangement with his disgraced predecessor but Mr Ford maintained: "There was no deal, period, under no circumstances." Even his most bitter critics later acknowledged that the pardon was an honest, necessary and politically brave act. Senator Edward Kennedy, the senior Democrat who opposed it, later said it was "an extraordinary act of courage that historians recognise was truly in the

national interest".

Mr Ford became vice-president in December 1973 after Spiro Agnew resigned following bribery allegations. Eight months later, he became president when Mr Nixon stepped aside rather than face a Senate trial on charges connected with the burglary of the Democratic National Committee headquarters. While president, he survived two assassination plots within 17 days, but still maintained a White House notable for its openness and informality.

Mr Ford was a high school football star from Grand Rapids, Michigan, who described himself as a "moderate in domestic affairs, an internationalist in foreign affairs and a conservative in fiscal policy".

As president, he seemed the epitome of an ordinary American. The day after his inauguration he made his own breakfast. But his everyman persona became fodder for ridicule. He was lampooned as a clumsy dolt and in June 1975 fell down the steps of Air Force One. Six months later he wiped out on the ski slopes. President Lyndon Johnson was widely quoted as having stated that Congressman Ford was "so dumb he can't walk and chew gum at the same time". In fact, the jibe was crueler and coarser – Johnson had actually said "fart and chew gum".

Bob Hope, who like Mr Ford was a golfing enthusiast, once joked: "It's not hard to find Jerry Ford on a golf course – you just follow the wounded."

The Vietnam War ended in ignominy for America during his presidency. After the fall of Saigon in April 1975, Mr Ford, an early supporter of the war who urged heavier bombing, said: "Today, America can regain the sense of pride that existed before Vietnam. But it cannot be achieved by re-fighting a war that is finished as far as America is concerned."

Damaged by an ailing economy and a fierce primary election challenge from Ronald Reagan, he lost the Oval Office to Jimmy Carter.

Betty Ford, who announced her husband's death, was a divorcee and dancer when they married 58 years ago. She admitted to alcoholism and addiction to prescription drugs while First Lady and later founded the Betty Ford Centre for alcohol treatment and drug rehabilitation.

She said yesterday that despite suffering a series of health scares in recent years, her husband lived a life filled by "God, family and his country". "My family joins me in sharing the difficult news that Gerald Ford, our beloved husband, father, grandfather and great grandfather has passed away," the former First Lady said in a statement.

ii. **Key moments in the life of Gerald Ford**

(*Telegraph*, 28 Dec., 2006)

Dec 6, 1973: Sworn into office as vice-president in the House of Representatives chamber using a Bible given to him by his oldest son, while his wife Betty looked on. He had refused Richard Nixon's insistence that he take the oath in the White House. Americans already believe he will soon become their president.

Aug 9, 1974: Nixon resigns and is replaced by Ford, who declares in his inauguration speech: "My fellow Americans, our long national nightmare is over. Our Constitution works. Our great Republic is a government of laws and not of men. Here, the people rule."

Sept 8, 1974: Grants a "full, free, and absolute pardon unto Richard Nixon for all offences against the United States which he has committed or may have committed or taken part in". Critics accuse him of striking a deal with his predecessor.

April 30, 1975: Americans are lifted from the roof of the US Embassy in Saigon by helicopters as the city falls to North Vietnamese troops. A week earlier, Ford had said that the Vietnam war "is finished as far as America is concerned".

June 1, 1975: Falls down the steps of Air Force One after slipping in the rain as he arrived in Austria. He says in a speech to his guests later in the day: "Thank you for your gracious welcome to Salzburg – and I am sorry that I tumbled in."

Sept, 5 1975: Lynette "Squeaky" Fromme, a member of the notorious Manson Family cult, is wrestled to the ground in Sacramento as she aims a loaded pistol at Ford as he leaves the Senator Hotel. Just 17 days later, Sara Jane Moore, a Left-wing zealot, fires two shots at Ford in San Francisco. She misses.

Nov 8, 1976: Jimmy Carter, the Democrat governor of Georgia, narrowly beats Ford to the presidency. Ford is damaged by a primary challenge from the conservative Ronald Reagan, a sluggish economy and his pardon of Mr Nixon.

Oct 4, 1982: Ford's wife, who had been an alcoholic and addicted to pain killers, opens the Betty Ford Centre in California for the treatment of drug and alcohol addiction.

iii. **Gerald Ford dies; Nixon's successor in '74 crisis was 93**

(*The New York Times*, 27 Dec., 2006)

Former President Gerald R. Ford, who was thrust into the presidency in 1974 in the wake of the Watergate scandal but who lost his own bid for election after pardoning President Richard M. Nixon, has died, according to a statement issued late last night by his wife, Betty Ford.

He was 93, making him the longest living former president, surpassing Ronald Reagan, who died in 2004, by just over a month.

The statement did not give a cause, place or time of death, but Mr. Ford, the 38th president, had been in and out of the hospital since January 2006 when he suffered pneumonia, most recently in October at the Eisenhower Medical Center in Rancho Mirage, Calif., for medical tests. He returned to his home in Rancho Mirage after five days of hospitalization.

"My family joins me in sharing the difficult news that Gerald Ford, our beloved husband, father, grandfather and great grandfather has passed away at 93 years of age," Mrs. Ford said in a statement issued from her husband's office in Rancho Mirage, also the location of the Betty Ford Center. "His life was filled with love of God, his family and his country."

President Bush praised Mr. Ford for his contributions to the nation "in an hour of national turmoil and vision," in a statement released early today from his ranch in Texas.

"With his quiet integrity, common sense, and kind instincts, President Ford helped heal our land and restore public confidence in the presidency," Mr. Bush said. "The American people will always admire Gerald Ford's devotion to duty, his personal character, and the honorable conduct of his administration."

Mr. Ford, who was the only person to lead the country without having been elected as president or vice president, occupied the White House for just 896 days – starting from a hastily arranged ceremony on Aug. 9, 1974, and ending with his defeat by Jimmy Carter in 1976. But they were pivotal days of national introspection, involving America's first definitive failure in a war and the first resignation of a president.

After a decade of division over Vietnam and two years of trauma over the Watergate scandals, Jerry Ford, as he called himself, radiated a soothing familiarity. He might have been the nice guy down the street suddenly put in charge of the nation, and if he seemed a bit predictable, he was also safe, reliable and reassuring. He placed no intolerable intellectual or psychological burdens on a weary land, and he lived out a modest philosophy. "The harder you work, the luckier you are," he said once in summarizing his career. "I worked like hell."

Gerald Rudolph Ford was born on July 14, 1913, in Omaha to Leslie Lynch King and Dorothy Ayer King. He rose to House minority leader in 1963 and served in the House until 1973, when Vice President Spiro T. Agnew resigned,

and President Nixon appointed Mr. Ford to succeed Mr. Agnew.

When Mr. Ford took the oath of president in 1974, the economy was in disarray, an energy shortage was worsening, allies were wondering how steadfast the United States might be as a partner and Mr. Nixon, having resigned rather than face impeachment for taking part in the Watergate cover-up, was flying to seclusion in San Clemente, Calif.

There was a collective sense of relief as Mr. Ford, in the most memorable line of his most noteworthy speech, declared that day, "Our long national nightmare is over."

Appendix B

i. 福特总统的传奇人生
(The Legendary Life of President Ford)

美国前总统杰拉尔·鲁道夫·福特于12月26日去世，享年93岁，是迄今美国历史上最长寿的总统。

鲁道夫·福特生性温厚谦逊、随和，以至于林登·约翰逊总统取笑当时还是议员的福特甚至“不能一边走路，一边嚼口香糖”。然而就是这样一位“好好先生”，一路福星高照，将美国从“水门事件”以及越南战争的阴影中解脱出来，成就了传奇一生。

ii. 史无前例：未经选举当上副总统和总统
(Unprecedented: President and Vice-president without Election)

1973年10月，副总统阿格纽因贪污丑闻辞职，福特因其良好形象和与尼克松总统多年的良好私交而被任命为副总统。八个月之后，尼克松因“水门事件”被迫辞职，福特直接就任美国第三十八任总统，福特成为美国历史上唯一一位未经选举而接连担任副总统和总统职务的人。

福特接手的是一个经过越战的烂摊子，这注定他的总统之路将荆棘密布，而尼克松犯下的错误，更是打碎了美国人对政府的最后一点信任。福特的首要任务就是恢复民众对白宫的信任。布什总统在追悼词中评价福特总统用其“不事张扬的正直、人所共有的感性和善良的本能”帮助美国人治愈了创伤，恢复了公众对国家的信心。

1975年4月，福特宣布越南战争结束，他带领美国人走过了越战结束后最困难、最混沌的日子。在内部最混乱的时候，重新团结国家，鼓舞人心，这是福特短暂总统的生涯的最大成就。

iii. **尴尬往事：摔下“空军一号”**
(Embarrassing: Falling down from Air Force One)

1975年六月一日，福特前往奥地利访问，那天刚下雨，他不慎从“空军一号”的楼梯上摔下。事后，他在演讲中自嘲说：“感谢你们的热情欢迎，抱歉我摔进了萨尔斯堡”。

除了人们的不理解，福特还面临着异常严峻的经济形势。福特虽殚精竭虑，但1975年1月失业率仍创下了33年来的最高纪录。全国各地响起一片谴责政府无能的抗议声。与此同时，柬埔寨和南越的亲美政权随着美军的撤出相继垮台，福特匆忙派兵到西贡帮助撤侨，千家万户的电视上出现了难民抱着直升机起落架仓皇逃命的画面。福特给美国民众留下了妥协退让、软弱无能的印象，有人甚至戏称他为“白宫里的童子军”

iv. **福星高照：躲过两次暗杀**
(Lucky: Surviving Two Assassinations)

福特虽然在任期内没有干出多少像样的政绩，但他温厚谦逊、随和的个性却让大多数人对他尊敬有加，称他是美国最老实的总统。命运女神也在眷顾着这位“好好先生”。在任期间，福特曾两次躲过暗杀，大难不死。1975年9月5日，福特夫妇参加一次集会时，一位红衣女郎掏枪对准他，但是因为枪支故障而没能射出子弹。十几天后，福特在观看一次展览时，一位当过联邦调查局密探的女士从提包里掏出左轮手枪，向大约40英尺外的福特射出一发子弹。不过因为看热闹的人破坏了杀手的视线，子弹偏离了几英尺。

v. **成也“水门”，败也“水门”：赦免尼克松丢总统宝座**
(Controversial: Pardoning Nixon and Losing the Presidency)

尼克松困水门事件无奈辞职，然而福特作出一个争议性很的决定，他宣布动用总统特权，无条件赦免涉及尼克松的所有罪行。对此，美国舆论一片哗然，咒骂其徇私枉法之声不绝于耳。也正因为这个原因，福特因此付出巨大的政治代价。2年后，他输掉了总统大选。福特在后来的总统选举中败给民主党总统假选人吉米·卡特，赦免尼克松被认为是导致失利的主要因素。然而当纷扰远去，尘埃落定后，美国人开始理解福特当时的良苦用心。许多曾极力批判这一赦免的人后来都不得不承认“这是极富勇气，真正为国家利益着想的举措”。

福特的一生既平凡而又不平凡，也许他的妻子蓓蒂对他的评价才最为恰当：“我的丈夫一生都充满对上帝、家人和祖国的爱。”