Editorial note

We recommend that references to the papers in these proceedings follow the format below:


Published by the University of Newcastle October 2015.

Callaghan, NSW

The copyright of each article remains with its author(s).

ISBN-10: 0-9941507-0-9

Table of Contents

Squib on Polish Yers: ............................................................................................ 4
An Overview of what we know, and how we got there
Brian Collins

The Present Perfect in Australian English narratives: ............................................ 29
Some preliminary sociolinguistic insights
Sophie Richard

Switch-reference and the subject in New Guinea: ............................................. 50
The case of Yagaria
Glenn Windschuttle

The Virtual Atoll Task: ....................................................................................... 82
A spatial language elicitation tool
Jonathon Lum and Jonathan Schlossberg

Epistemic modulations and speakers stance in Cantonese conversations .... 104
Winnie Chor

Predatory chatlog messages as forensic evidence in court: ............................. 131
A comparison of two different procedures for estimating the weight of evidence
Shunichi Ishihara

Loan phonology in Murrinhpatha ................................................................. 153
John Mansfield

Where’s wh-? Prosodic disambiguation of wh-scope ................................... 173
Iain Giblin and Sam Steddy
Squib on Polish Yers: An Overview of what we know, and how we got there

Brian Collins
University of Utah
u0949515@utah.edu

Abstract: This squib puts various explanations of Polish vowel-zero alternations to the test, compares them, and critically analyzes some of the assumptions the explanations use. Historically, these alternating vowels (known as Yers) have presented serious opacity problems as they occur with high frequency in Polish, even in some loan words. A repair to the OT approach is developed Anchor-Yer-L, that accurately predicts which Yer will delete. This paper also reanalyzes the productivity of Polish vowel-zero alternations using conversations in the Polish lexical data.

Keywords: Yer, Optimality Theory, Sonority Sequencing, Sonority Reversal, Opacity

1. Introduction
Many Slavic languages exhibit vowel-zero alternations in their phonologies. The vowels, which synchronically alternate with zero, are both synchronically and diachronically called Yers in Slavic Linguistics. In historic contexts (see section 2) I shall disambiguate between the synchronic Yer, or the sound that occurs in contemporary Slavic languages, which alternates with zero, and the diachronic Yer, or the short vowel(s) the synchronic Yer historically was.

A synchronic Yer occurs as a vowel (in Polish, as a full vowel ‘e’ /ɛ/ or /e/) when it is the nucleus of the final syllable of a word, but then fails to surface in
other contexts (data from a combination of my own personal knowledge and Szpyra (1992)):

(1)

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>sweter [sfɛters]</td>
<td>swetra [sfetra]</td>
</tr>
<tr>
<td>pies [pjɛs]</td>
<td>psa [psa]</td>
</tr>
<tr>
<td>susel [susew]</td>
<td>susla [suswa]</td>
</tr>
<tr>
<td>Marek [marek]</td>
<td>marka [marka]</td>
</tr>
<tr>
<td>dizel [dizel]</td>
<td>dizla [dizla]</td>
</tr>
</tbody>
</table>

If the data were as simple as this, the problem could be solved with a simple epenthesis explanation to repair a coda consonant cluster. However, we are then faced with other cases where no Yer ever surfaces and a coda consonant cluster remains:

(2)

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>menawr [menavr]</td>
<td>menawru [menavru]</td>
</tr>
<tr>
<td>Piotr [pjɔtr]</td>
<td>Piotra [pjɔtra]</td>
</tr>
<tr>
<td>wiatr [vjatr]</td>
<td>wiatru [vjatru]</td>
</tr>
<tr>
<td>pieśn [pjɛnɛn]</td>
<td>pieśni [pjɛnɛni]</td>
</tr>
</tbody>
</table>

And yet further cases where an [ɛ] vowel remains despite the addition of another syllable:

(3)

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>rower [rówɛr]</td>
<td>roweru [rówɛru]</td>
</tr>
<tr>
<td>bies [bjɛs]</td>
<td>biesa [bjɛsa]</td>
</tr>
</tbody>
</table>
In some Slavic languages, the number of roots where vowel-zero alternations occur synchronically is limited, but they do still exist in many suffixes. These root alternations are still of particular interest in Polish however, because they occur throughout the lexicon. There are many 20th century loan words such as *sweter* ['sweater'] (Gussman, 1980) and *diesel* ['diesel'] (Rubach, 1984) which have vowel-zero alternations. Jarosz (2005) explicitly states that all three paradigms above in (1) (2) and (3) (alternating vowels, complex coda, and non-alternating vowels) are fully productive; the judgments of native Polish speakers quoted on the SJP lexical database seem more mixed however (see section 6.2).

Szpyra (1992) estimates that Yers occur in 63% of the Polish lexicon that would otherwise end in a sonority reversing Cr (any consonant and /r/) coda. While this shows how common Yers are in Polish, it still leaves a large proportion of the lexicon simply ending in a sonority reversing Cr coda.

Early approaches such as Gussman (1980) and Rubach (1984) treated Yers as abstract vowels, which never surface (either deleting or becoming [e]). Later approaches such as Szpyra (1992) have treated Yers as underlyingly underspecified segments in a segmental approach, or within the framework of Optimality Theory, such as Jarosz (2005b).

A modification of a solution by Jarosz (2005b), which accounts for all of the vowel-zero alternations, will be presented in section 5.

### 2. Historical Overview

1 The official Polish lexical database unfortunately does not have the first attested date of words. Logically *sweter* and *diesel* must have been borrowed during the 20th century, given that the word *sweter* was borrowed from the English word ‘sweater’ which was not used to describe clothes until 1882 (Harper, 2015); diesel was first used to describe the fuel source in English 1892 (Harper, 2015), and presented at the Paris Exposition in 1900. The first Polish etymological dictionary compiled by Brückner (1927) lacked both words. The diachronic shift (see section 2) which created the synchronic vowel-zero alternations is believed to have begun the 10th century. It is generally taken for granted that this shift has completely eliminated any phonetic trace of the Late Common Slavic short or lax vowels from Polish.
In Proto-Slavic these alternating segments were originally two short high vowels *
[ū] and *
[į], or possibly two lax vowels as characterized by Lunt (2001). They were vocalized in all environments and never deleted. In the older attested Slavic languages, such as Old Church Slavonic, the front Yer was represented with the letter ь while the back Yer was represented with the similar letter ъ. These vowels occurred in many more environments than synchronic Yers do in contemporary Slavic languages.

Late Common Slavic (LCS), and Old Church Slavonic, the oldest attested Slavic language, did not allow coda consonants and had many more phonotactic constraints than contemporary Slavic languages like Polish. Yers occurred in native words, and were also used to repair illegal consonant clusters and form open syllables (Data from Lunt, 2001, p. 34; transcriptions are my own):

(4)

<table>
<thead>
<tr>
<th>Greek</th>
<th>Old Church Slavonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>/pavlos/</td>
<td>/pavilŭ/</td>
</tr>
</tbody>
</table>

Before LCS diverged into the different Slavic language branches, a phonological process began, which was characterized by Schenker (1993) as “the weakening of the Yers”, and by others (Townsend and Janda (1996)) as “the Fall of the Yers”. “Weak Yers” deleted, while “strong Yers” became full vowels: “the jers were weak in word final position, strong before a weak jer, and weak before a strong jer or any other vowel.” Data from Townsend and Janda (1996, p. 73) (transcriptions my own):

(5)

<table>
<thead>
<tr>
<th>Late Common Slavic</th>
<th>Polish:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) sŭnŭ sŭna</td>
<td>syn [śn] sna</td>
</tr>
<tr>
<td>son (NOM) son-GEN</td>
<td></td>
</tr>
<tr>
<td>b) zamŭkŭ zamęk</td>
<td>castle (NOM)</td>
</tr>
</tbody>
</table>
From another point of view, Yer deletion alternated starting from the right edge of the word. This can be seen in d) where the rightmost Yer deleted, but the central Yer remained. The central Yer was in a strong position as it was to the left of a weak Yer, by implication, because the initial Yer was to the left of a strong Yer it deleted.

As Townsend and Janda (1996, p. 73) note however, as in d) “in positions where a multiple vowel/zero alternation was motivated… one form has been leveled through the paradigm in most Slavic languages”. In other words, multiple vowel zero alternations do not occur synchronically in most Slavic languages (including Polish).

Importantly, from a synchronic perspective, Yers were historically epenthesised to repair phonotactic violations, in addition to occurring in words inherited from Proto-Slavic. Also important for an analysis in sections 4 and 5, Yers deleted from the right edge of the word predictably.

3. Pertinent Phonological Facts

3.1 Polish Stress

Unlike most Slavic languages such as Russian Ukrainian, which have phonemic stress, or Serbo-Croatian, which has phonemic tonal contours (see Inkalas & Zec (1988) for a discussion on Serbo-Croatian), Polish stress is predictably
penultimate, with the exception of a few Latinate loan words like *fizyka* ['fizika], and *matematika* [matɛ’matika] which have antepenultimate stress.

Some authors, such as Jarosz (2005a), have analyzed the Polish Yer with regards to stress. This seems intuitive given that Yers most commonly delete from the position where they would otherwise be penultimate. When a Yer deletes, it therefore commonly maintains the number of syllables in a word because the suffix being added has only one syllable:

(6)

[ˈsfɛtɛr]→[ˈsfɛtra] not *[sfɛˈtera]

Jarosz (2005a) claimed that maintaining the syllable count, and therefore maintaining the stress on the same vowel throughout the declensional paradigm was the reason for Yer deletion.

This would explain the situations in (1) and (2), where either a vowel alternates with zero, or a coda consonant cluster is permitted. Polish would delete [e] in *dizel* ['dizɛl] or *sweter* ['sfɛtɛr] when another vowel is added, in order to keep the stress on the initial syllable.

(7)

Nominative   Genitive

[ˈdizɛl]       [ˈdizɛla] *[di´zela]

[ˈsfɛtɛr]       [ˈsfɛtra] *[sfɛˈtera]

By maintaining the number of syllables, it keeps the penultimate stress on the same syllable rather than shifting it to a new penultimate position when another syllable is added. This explanation however fails to account for the data in paradigm of (3) reiterated in (8):

(8)

Nominative   Genitive

[ˈkrɛtɛr]       [kra´teru] *[´kratru]
Polish [ɛ] should always delete if we postulate that it deletes when one syllable is added, in order to maintain the syllable count, and therefore maintain the stress on the same syllable throughout the paradigm.

If Polish suffixes conveniently only ever had one syllable, this approach could describe Yer deletion with just one stipulation (see next section). However there are Polish affixes with multiple syllables, such as –ami [ami], the instrumental plural marker:

(9)

sweter  [ˈsfɛtɛr]  ‘sweater(-NOM.SG)’
swetr-ami  [sfɛˈtrami]  ‘sweater-INSTR.PL’

It is thus not only penultimate Yers that delete but any Yer that does not occur in the last syllable of a word, regardless of the number of syllables the suffix has.

3.2 Lower

Another important fact regarding vowel-zero alternations is that, in situations with multiple synchronic Yers, only the rightmost Yer deletes. This is similar to the diachronic change in which the Yers in word-final position were “weak” and deleted, while the preceding "strong" Yers were fully realized. Formally this rule is called Lower in the literature (see Szpyra (1992), Jarosz (2005), Gussman (1980) Rubach (1984; 1986) for further discussions).

This is observable with the Polish diminutive suffix ek(a/o), which has a Yer. When it is affixed to a word that already has a Yer in it, the Yer remains (Yers in bold and underlined):

(10)

sweter  [sfɛtɛr]  ‘sweater (NOM.SG)’
swetr-a  [sfɛtra]  ‘sweater-GEN.SG’
swetɛr-ek  [sfɛtɛrɛk]  ‘sweater-DIM’
swetɛr-ka  [sfɛtɛrka]  ‘sweater-DIM-GEN.SG’
Like we have seen diachronically, the synchronic Yer remains if there is another Yer to the right of it. Szpyra (1992) wrote this rule essentially as follows:

(11)

Lower:

\[ \text{Yer} \rightarrow e/\_CYer \] (a Yer surfaces if another Yer occurs after it.)

4. Approaches to Analyzing the Yer

4.1 Epenthesis Approach

Despite the obvious challenges to an epenthesis approach, some authors Gorecka (1988), Czaykowska-Higgens (1988), and Pietrowski (1988), proposed an epenthesis analysis to explain vowel-zero alternations in Polish. As mentioned in the introduction, in the majority of cases involving vowel zero-alternations a vowel breaks up a sonority reversing coda.

The problem however, is that Polish does permit sonority-reversing coda in a large number of words. As Szpyra (1992) indicated, word final Cr are only broken by Yers 63% of the time in the lexicon.

Another major problem noted in Szpyra (1992) with an epenthesis analysis, is that there are many minimal pairs in which a Yer breaks up the final consonant cluster in a noun in certain declensional paradigms, where -ł (IPA /w/) occurs as the final consonant before the gender suffix (-a for feminine nouns and –o for neuter nouns); the masculine past tense of a verb in takes a suffix –ł [w] too, only, unlike in the noun, no Yer surfaces:

(12)

a)  

miotł-a [m̩ɔtɛw]   ‘broom-NOM.FEM’

mioteł [m̩ɔtɛw]    ‘broom (GEN.PL)’
mióił [miɔtw] ‘sweep-PAST.MASC’

b) 
piekł-o [pʲɛkɛwɔ] ‘hell-NOM.NEUT’
piekel [pʲɛkɛw] ‘hell (GEN.PL)’
piek-ł [pʲɛkw] ‘bake-PAST.MASC/he baked’

Note the lack of a Yer in the nominative case in paradigms a) and b) when there is a gender suffix –a or –o. This illustrates the existence minimal pairs between the same roots, where nouns have a Yer, and verbs do not. As Szpyra (1997) noted, an epenthesis analysis would not predict minimal pairs with vowel-zero alternations, because it assumes the alternations exist for phonological reasons, and that the Yer does not exist in the underlying form.

Szpyra (1992) also notes examples of minimal pairs between nouns, where again the only difference is the presence of a Yer (data from Szpyra (1992)):

(13)

a) 
bark [bark] ‘shoulder (NOM.SG)’
bark-i [barki] ‘shoulder-PL’

b) 
bar-ek [barek] ‘bar-DIM’
bar-k-i [barki] ‘bar-DIM-PL’

‘Shoulder-PL’ and ‘bar-DIM-PL’ (both pronounced [barki]) are homophones, but their singular forms are minimal pairs because of the presence of a Yer. If Yers existed as an example of phonological epenthesis to break up marked sequences of sounds, then minimal pairs between a Yer surfacing, and a Yer not surfacing, would of course be unexpected.
The epenthesis analysis however does capture one generalization which other analysis (see next section) do not: the position of Yers is mostly predictable. They occur primarily in the final syllable of a root between an obstruent and an unsyllabic sonorant: [l] [w] [r] [n] [m].

4.2 Abstract Approaches

Some of the earliest approaches (Gussman, 1980; Rubach 1984, 1986) to analyzing the Yer in Polish involved an abstract analysis: essentially, they assumed underlying “segments that appear nowhere on the surface”, as Hayes (2011) put it.

For simplicity, I will assume there is one underlying Yer in these analyses as Hayes (2011) and Jarosz (2005b) do. The early abstract analysis essentially works like this: the Yer is underlingly a vowel Polish does not have, most commonly described as /ɪ/; the Yer either deletes, or lowers to [ɛ] on the surface (rule 2 from Hayes (2011), rule 1 modified from Lower in Szpyra (1992)):

(14)

Rule 1)
Lower:
ɛ/_C1

Rule 2)
Yer Lowering:

---

2 The initial by Gussman (1980) and Rubach (1984) actually assumed multiple underlying abstract segments that would either surface or delete. Yers which seemed to trigger palatalizations, were explicitly distinguished from Yers which did not. This is observable in the vowel-zero alternations for [pjɛs]–[psɪ] alternations; the word has a palatalized /pj/ when the vowel surfaces, but does not when it deletes. In the [sfɛtɛr~sfɛtɛn] alternations however, the [ɛ] is never palatal even when the Yer surfaces.
Rule 3)

Yer Deletion:

/1/ → ∅

If the rules are applied in that order, a Yer in an environment to which Lower or Yer Lowering applies will become [ɛ] on the surface, elsewhere it will delete:

(15)

<table>
<thead>
<tr>
<th></th>
<th>sfetr</th>
<th>sfetrami</th>
<th>sfetrk</th>
<th>menavr</th>
<th>røveru</th>
<th>---</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>---</td>
<td>---</td>
<td>sfetrk</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Yer Lowering</td>
<td>sfetr</td>
<td>---</td>
<td>sfetrk</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Yer Deletion</td>
<td>---</td>
<td>sfetrami</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Surface Forms</td>
<td>sfetr</td>
<td>sfetrami</td>
<td>sfetrk</td>
<td>menavr</td>
<td>røveru</td>
<td>---</td>
</tr>
</tbody>
</table>

There is no doubt that this analysis accounts for all the data. Yer lowering does not apply to sfetrami because the Yer is not at the edge of the word. Yer deletion does not apply after Yers have been lowered, and nothing is epenthesized in [menavr], the [ɛ] in [røveru]. This analysis is not without critics however. As Hayes (2011) puts it

To take the abstract analysis seriously, we interpret it to mean that when Polish-learning children encounter the data conundrum that we have just encountered, they respond to it by expanding their underlying vowel systems to include jers. In a sense, Polish children are claimed to invent the jers intuitively as a means of handling the data, just as clever phonologists have done in their consciously guided analytical work. This claim is potentially problematic. One might be wary, for instance, of a theory in which the right answer must come from a lively act of the imagination – for what happens to the language acquirer who through bad luck or inferior skill happens never to arrive at the right inspiration?” (p. 242).

The obvious weaknesses of an abstract approach notwithstanding, there are not many alternatives to an abstract analysis of Yers. Other rules, where the environment
triggering an alternation has been diachronically deleted, such as English f–v alternations (e.g. knife–knives), have been shown to be semi-productive synchronically (see 6.1).

Even Yer epenthesis, the primary alternative, can be traced back to a diachronic process, where Yers were epenthesized in LCS words to repair phonotactic violations. Modern Polish itself has much less restricted phonotactics than LCS, and allows for coda clusters and violations in sonority sequencing.

4.3 A Reanalysis of the Underlying Properties of Yers

Szpyra (1992) thought it was possible to determine the phonological properties of Yers in order to best formulate a rule describing their deletion.

Bethin (1987) and Rubach (1986) and Szpyra (1992) claimed that Yers fail to undergo syllabification, which is what mainly distinguishes them from other vowels. Rubach & Booij (1990) and Bethin (1989) used an allomorphic rule to demonstrate this. In Polish, comparatives are formed by adding -sz [ʂ] or -ejsz [eiʂ] to the end of an adjective’s root; for example ciekawy [tɕɛkavɨ] ‘interesting’ and ciekawszy [tɕɛkavʂɨ] ‘more interesting’; ciepły [tɕɛpɨ] ‘warm’ and cieplejsze [tɕeplɛjsɨ] ‘warmer’. Szpyra (1992) formulized a rule as follows to describe this process (p. 286):

(16) $\emptyset \rightarrow [ei]/ *C_\_ [s]\$ epenthesize an [ei] before [s], if [s] occurs after an unsyllabified consonant.

Some Polish adjectives and adverbs also have Yers; for example pewien [pɛvʲɛn] ‘certain (adv)’ has the adjectival form pewny [pɛvnɨ]. One theory-internal assumption is that allomorphic rules generally precede phonological rules. If Yers are syllabified like normal vowels in the UR, and the allomorphic rule occur before Yer deletion, Szpyra (1992) argues it should result in a comparative form of *pewnszy but it does not; it instead results in pewniejszy:

If the adjective pewien has two syllables in the UR and the allomorphic rule was ordered before Yer deletion it would be:
If, however, we assume that it were not syllabified in the UR, the allomorphic rule would ignore the Yer, and behave as if it were underlyingly /pɛvɛn/ resulting in *pewnījszy*. This argument relies on an assumption that the allomorphic rule applies because of an unsyllabified consonant, and also that Yer deletion must occur before comparative allomorphy.

Szpyra (1997) proposes a solution to the problems, suggesting that Yers might be neither consonantal nor vocalic in the UR:

In other words, the proposal is to recognize underlying segments that are neither vowels nor consonants, i.e. units that are maximally underspecified even with respect to major class features. Such segments do not undergo Syllabification, since without the feature [consonantal] they cannot be assigned any place in the syllable… This means that a melodic element, including an empty root node, when found outside a syllable, will block Syllabification. Thus, Yers viewed as empty root nodes will escape Syllabification and at the same time be nontransparent with regard to it—exactly the characteristics required for these elements… The rule that vocalizes Yers can be seen as supplying the feature [−consonantal] to empty root nodes. (p. 299)

Essentially, Szpyra proposes decreasing the assumed features to explain both the failure of Yers to syllabify and their predictability. This approach is still abstract in nature, insofar as it assumes speakers have an underlying unit in vowel-zero alternating words, which is neither a vowel nor a consonant.

One problem with Szpyra (1992), and others’ conclusion that Yers are immune to syllabification is that a simple reordering of the rules creates the proper surface form:
It may be theoretically more viable to delete a Yer after the allomorphic rule applied, but one must consider all possibilities, rules can simply be ordered differently. One of the primary reasons\(^3\) Szpyra (1992), Rubach (1986), and Bethin (1989) suggest Yers are invisible to, or block, syllabification is a theory-internal reason; allophonic rules apply after allomorphic rules.

It is clear Polish does allow for consonant clusters of stops and fricatives (even the same fricative) before [ʂ] in the comparative forms of words like \(\text{droższy}\) [\(\text{drɔʂʂɨ}\)] ‘more expensive’ and \(\text{twardszy}\) [\(\text{twardɨʂɨ}\)]. The only time we see the allomorph -ejsz is after sonority reversing roots for adjectival endings such as [-vn] [-pl] [-dn]. This rule would also capture the allomorphy and does not assume any consonants are unsyllabified:

\[
\emptyset \rightarrow \text{ej} / \text{C[+sonorant -syllabic]-ʂ epenthesize [ej]} \text{ after a consonant cluster}
\]  

\[
\text{ending in a sonorant.}
\]

There simply are no instances of [ej] being epenthesized except for in cases where the consonant cluster ends in an unsyllabic sonorant, namely [l] or [n].

\[4.5 \text{An OT Approach:}\]

\(^3\) There were other allomorphic arguments involving Polish imperative forms, and alternations with the high back vowels. Results from Wug Tests (see Sanders (2003)) have indicated that Polish back vowel alternations are no longer synchronically productive. The vowel alternations themselves are beyond the scope of this paper, but they are the result of the LCS morphological long vowel alternations between [o:]~[o] with [o:] eventually shifting to [u] in Polish. Unlike with Yers, there are few loan words that undergo the alternation. The only productive alternations are morphologically motivated.
Jarosz (2005b) approached vowel-zero alternations in Polish through the scope of Optimality Theory. Jarosz (2005b), like in the abstract analysis in 4.2, assumes that the value of a Yer is underlingly [i] but undergoes lowering. By ranking */i>*IDENT[HIG], MaxV, Jarosz was able to explain alternations in words with complex codas such as the sweter/swetry case (recall Example 7):

(20)

<table>
<thead>
<tr>
<th></th>
<th>*/i</th>
<th>*COMPLXCD</th>
<th>ID[H]</th>
<th>Max-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>sfɛtr∅</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfɛterr</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
| sfɛtr | *! | | | *
| sfɛ’tr | *! | | | | |

If a Yer is not needed to repair complex coda, it deletes:

(21)

<table>
<thead>
<tr>
<th></th>
<th>*/i</th>
<th>*CMPLXCD</th>
<th>ID[H]</th>
<th>Max-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>sfɛtr-a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfɛtera</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>sfɛtra</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>sfɛ’tra</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Jarosz explains complex codas by ranking Dep-V above *Complex Coda:

(22)

<table>
<thead>
<tr>
<th></th>
<th>*/i</th>
<th>DEP-V</th>
<th>*CMPLXCD</th>
<th>ID[H]</th>
<th>Max-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>pjoter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pjotr</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pjotr</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
This analysis explains why in clusters such as CCC a Yer will only ever vocalize between the second and third consonants (CCVC). If it vocalizes anywhere else, it violates ID[HI] and fails to break up a complex coda.

Jarosz (2005) explains multiple Yers voicing (Lower) by applying MAX-V a second time, creating a new constraint, MAX-V2. This is a prohibition on deleting a second consecutive vowel. By then ranking MAX-V2 above MAX-V1:

\[(23)\]

<table>
<thead>
<tr>
<th>sfetırk-∅ sfetırk-a</th>
<th>*CMPLXCD</th>
<th>CI</th>
<th>MAX-V2</th>
<th>ID[HI]</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>sfeterk, sfeterka</td>
<td>!</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>sfetrek sfetrek</td>
<td>!</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfetrek sfetrek</td>
<td>!</td>
<td>*</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfetırk sfetırk</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfetırk sfetırk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thus, under this analysis, the Lower rule becomes obsolete. A second violation of MAX-V is not permissible, which accounts for Yers being arbitrarily vocalized. Contextual identity of the morphemes prohibits a Yer from surfacing in the wrong place. This constraint could easily be replaced with a more phonological constraint. Which Yer is deleted is phonologically predictable regardless of the morpheme boundary, as discussed below in section 5.

5: Anchor-Yer-L

Specifically for the OT analysis by Jarosz (2005), boundaries between morphemes are brought into the description of where Yers delete. This unnecessarily
complicates the analysis. This analysis could be simplified by focusing on a single boundary, the word boundary. If there are multiple Yers, and a Yer deletes, it is always the rightmost Yer from the word boundary.

This ranking could fit into the table, but would have to be below MaxV2. Otherwise Yers would be deleted where they should not be:

(24)

<table>
<thead>
<tr>
<th>sfetrık-Ø sfetrık-a</th>
<th>MAX-V2</th>
<th>Anchor-Yer-L</th>
<th>ID-HI</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>sfetrık-Ø sfetrık-a</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sfeterek-Ø sfetrek-a</td>
<td></td>
<td>*!</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>sfeterek-Ø sfetrek-a</td>
<td></td>
<td></td>
<td></td>
<td>**<em>↓</em></td>
</tr>
<tr>
<td>sfeterek-Ø sfetrek-a</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>sfetrek-Ø sfetrk-a</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>sfetrek-Ø sfetrk-a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With this new constraint, instances of Yers deleting in a position closer to the left, and thus leaving the Yer closer to the right, are harmonically bounded. Morpheme boundaries are no longer directly relevant, except for the word edge. The rightmost Yer deletes instead of the Yer further to the left not because of...
morpheme identities, but rather because it is simply further from the edge of the word.

This synchronic explanation also describes the diachronic process. In section 2, diachronically, the rightmost Yer was always in a weak position, and always deleted. To some, this may seem as a weaker analysis because it corresponds so well to the diachronic changes. As Hayes (2011) put it when criticizing the abstract analysis in 4.2, “these historical facts cannot be made to bear on the synchronic analysis of Polish, because Polish children learning the complex phonology of their language have no access to historical information” (p. 243).

In one important way, the synchronic solution does differ from the diachronic change: in the diachronic change Yers to the left of weak Yers were strong, while Yers to the right of strong Yers were weak; Yer alternations between weak and strong happened in a strong-weak-strong-weak pattern, whereas in this synchronic approach only the rightmost Yer is weak or strong-strong-strong-weak. Synchronously there is no evidence for Yer alternation. In situations with three Yers, for example only the rightmost Yer deletes. Polish allows for diminutive stacking:

(25)

<table>
<thead>
<tr>
<th>Nom</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>sweter-ecz-ek [sfeteretśek]</td>
<td>sweter-ecz-k-a [sfeteretśka]</td>
</tr>
<tr>
<td>sweater-dim-dim</td>
<td>sweater-dim-dim-GEN</td>
</tr>
</tbody>
</table>
‘sweater (NOM)’ ‘sweater-GEN’

/sfɛtər-ŋtʂ-ik-a/ surfaces as swetereczka [sfɛtɛɾɛtʂka] not *swetereczka *[sfɛtɛɾɛtʂka] as the purely diachronic rule would predict.

Many of the aspects of Yers predictability are also accounted for in this OT approach. Complex coda are permitted as in (22) because DepV is ranked above complex coda, but *ɪ is ranked above that. Yers primarily occur in environments where a word ends in a complex coda, and by addressing both complex codas and the abstract vowel *ɪ in rule ordering, we have the predictions of the epenthesis approach, but also the operational solution of the abstract approach. The final question: “which Yer deletes in situations where multiple Yers surface?” has been addressed by anchoring Yers to the left edge of a word, rather than by addressing the problem in morphology.

6. Reanalyzing the Productivity of Yers

6.1 Diachronically motivated partial productivity

Berko (1958) analyzed the English plural [f]~[v] alternation through Wug Tests. As with Polish Yers, English [f]~[v] alternation frequently happens, but then there are also many instances where [f]~[v] alternation does not occur. This alternation was once\(^4\) fully productive. Many English nouns which end in [f] voice the [f] in the plural to [v]:

\(^4\) In early stages of Middle English, many fricatives pairs /θ/~/v/ /θ/~/ð/ /s/~/z/ were in allophonic variation. The fricatives were voiced intervocally, and words like ‘leaves’ had one extra syllable [leːvəz] (see Kurath (1956) for a detailed discussion). At some point word final schwas began to drop, and thus, as with Yers, we now have a situation where the environment triggering phonological alternations has been diachronically erased.
(26)
Singular  Plural
knife [najf]    knives [naivz]
wife [wajf]    wives [wajvz]
leaf [lif]     leaves [livz]
wolf [wælf]    wolves [wælvz]

Then there are numerous nouns, which clearly, at least by my judgments, do not alternate between [f] and [v]:

(27)
gulf [gælf]    gulfs [gælfs] *gulves
whiff [wɪf]    whiff[s] [wɪfs] *whives
reef [ɹɪf]     reefs [ɹɪfs] *reeses

There are also nouns I feel much less certain about, even as a native speaker of English and (even with the orthography only giving one possibility in some cases):

(28)
life [læjf]     lives [læjvz] or [læjfs]?
mastiff [mæstɪf] mastiffs [mæstɪvz] or [mæstɪfs]?
scarf [skæɹf]  scarfs [skæɹfs] or scarves [skæɹvz]
dwarf [dwarf]  dwarfs [dwarfs] or dwarves [dwarvz]?

The scarf~scarves/scarfs and dwarf~dwarves/dwarfs even has two different possibilities acknowledged by English orthography. Wug tests have revealed that there is disagreement by native adult English speakers as to whether it applies to nonsense words, and that this disagreement changes with age:
It must be noted that although 42 % of the adults gave *heaves as the plural of this item, employing what would amount to a morphophonemic change along the lines of: knife: knives; hoof: hooves, only three children out of a total of 89 answering this item said ‘heaves’; 9, or 10 % added nothing, and an additional four formed the plural with the wrong allomorph, i.e. they said /hıyfəz/, treating the /-f/ as if it belonged to the sibilant-affricate series…. The proportion of children's right answers suddenly drops when we come to the form *tass. As table 3 shows, 91 % of these children when given the form glass could produce the form glasses. When given the form *tass, a new word patterned after glass, only 36 % could supply the form Classes. The picture becomes progressively worse with the other words ending in sibilants or affricates, and by the time we reach the form *niz, only 28 % answered correctly. (Berko, 1958, p. 162).

There were differences in applying allomorphic rules for different age groups. Berko (1958)’s experiment showed age gradation: preschoolers did worse than first graders for all examples.

Taking Hayes (2011)’s question mentioned in section 4.2 seriously “what happens to the language acquirer who through bad luck or inferior skill happens never to arrive at the right inspiration?” perhaps they never generalize a rule, or memorize allomorphic rules such as [əz] for [z] after a sibilant as exceptions. It sounds entirely plausible for some adults to say “three tass” (for a new word “tass”) instead of “three tasses”.

It also seems that the majority of English speakers never generalize a rule for f~v plurals. Only 42% of speakers apply the rule productively. What happens to the 58% of speakers who never generalize a rule? In a simplistic prediction (where we assume speakers either apply, or do not apply a rule with total consistency), we might assume they are the 58% producing [hıfs] instead of [hıvz]. The question then is, is there this same speaker variation that happens with English f~v alternations as there is for Polish vowel-zero alternations? The answer seems to be yes.

6.2 Polish Yer Productivity
While initial findings have not been formally tested through Wug Tests\(^5\), delving into the Polish corpus reveals conversations between native Polish speakers about vowel-zero alternations in loanwords, and confusion over them (data from SJP; translations by me):

**Question**: Czy poprawne są obie wersje: **swetr** i **sweter**?

‘Are both forms correct? swetr and sweter?’

**Answer**: Poprawna jest tylko forma sweter - wystarczy zajrzeć do któregokolwiek słownika, aby się o tym przekonać. Dłuższe wyjaśnienie błędu daje np. Słownik wyrazów kłopotliwych PWN.

‘The correct form is only sweter. Simply look at any dictionary. The longer explanation of the mistake provides an example: troublesome words dictionary PWN’

In this exchange, the first speaker seems to feel both swetr and sweter are correct for ‘sweater’, while the second speaker advises the first to read a dictionary to see that only sweter is correct. In some exchanges over this word pair, answerers provided not only completely incorrect, but descriptively peculiar explanations for why sweter is not swetr:

**Question**: Dzień dobry Państwu, Czy mogliby państwo wytłumaczyć mi nielogiczność językową w słowach: wiatr, ale nie wiater, sweter, ale nie swetr? Z góry dziękuję.

‘Good day to you all. Would it be possible for you to explain the illogical paradigms of these words: wiatr but not wiater, sweter but not swetr. Thanks in advance’

**Answer**: Wiater to forma gwarowa, sweter pochodzi z angielskiego sweater. Nie ma więc żadnej nielogiczności.

‘Wiater is a slang form, sweter comes from the English sweater. Thus, there is nothing illogical about it.’

---

\(^5\) To my knowledge, there have been no tests on the productivity of vowel-zero alternations in Polish. Gouskova & Becker (2013) performed tests to see if Russian speakers accept Yer deletion in nonce words. The grammaticality judgments showed that Russian sometimes accept novel vowel-zero alternations in polysyllabic nonce words, but judge the alternations in monosyllabic nonce words to be ungrammatical. This however did not test whether Russian speakers produce novel vowel-zero alternations, and assumed they are suppletive in nature. This may be true for Russian, which does not have vowel-zero alternations in loanwords.
This provides an interesting insight into how native Polish speakers perceive vowel-zero alternations, in a way that has not been revealed before. In the second exchange, the answerer seems to believe that *sweter* has a vowel-zero alternation because it is loanword from English, and the answerer also stated that an alternating vowel was possible for *wiatr* ‘wind’ but that this is ‘slang’. There do seem to be accepted standard forms, similar to the way *[lajvz]* is more standard than *[lajfs]* for ‘lives’ in English (at least per the orthography), but there is disagreement between native speakers, and some seem to be confused. This is very unlike the way the literature tends to portray these alternations. Jarosz (2005) for example, characterizes all three paradigms from the introduction as “fully productive” (p 1).

7. Conclusions & Further Directions

Assuming an underlying abstract segment, an OT approach can handle the data much like the rule-ordering approach is able to. While there are obvious weaknesses to an approach that operates with one segment never occurring on the surface, the alternatives do not seem to make the correct predictions. The OT approach further accounts for the predictability of Yers by tying their surfacing to a violation of complex coda as in (22).

As demonstrated in (12) and (13) (section 4.1) there are minimal pairs between words with vowel-zero alternations, and words without. An epenthesis solution, like the abstract solution, is also inherently linked to diachronic changes in Polish phonotactics. Synchronic Polish permits the coda clusters that an epenthesis approach fixes; diachronically, as shown in (4), LCS used Yers to repair consonant clusters and even create open syllables.

In the context of Opacity, there should be further verification for the productivity of Yers in Polish. They do occur in loanwords, but it seems that vowel zero alternations are only partially productive.

The interaction between vowel-zero alternations and the suffix –ek(a/o) is also of further interest. Some loanwords, namely *rober* [ɾɔˈber] ‘rubber’ which have vowel-zero alternations were listed with the diminutive suffix (e.g. *roberek*) as a ‘synonym’
to the form without a diminutive suffix in the SJP. From data already discussed, the diminutive suffix has an effect on Yers insofar as it preserves them, either through my constraint Anchor-Yer-L, or through the rule Lower. For children acquiring the language, it is entirely plausible that the suffix could have an effect on whether a the learner perceives a vowel [ɛ] in the final syllable of the root; if the learner encounters swetry first instead of sweterek or even sweterki, will this affect how the speaker conceptualizes the vowel, or whether the speaker even believes there is an underlying vowel in the root?

References


The Present Perfect in Australian English narratives:  
Some preliminary sociolinguistic insights

Sophie Richard  
University of Western Australia  
sophie.richard@research.uwa.edu.au

The Present Perfect has been documented in Australian English (AusE) to encode narrative progression (Engel & Ritz 2000; Ritz & Engel 2008). However, the sociolinguistic constraints on the Narrative Present Perfect (NarrPP) in AusE stories have yet to be investigated. In this paper I report on an exploratory study of a sub-sample of performed (Labovian) narratives produced by West Australian male and female speakers, aged 36 to 56, of professional and non-professional occupational backgrounds. The data stem from Richard’s (under construction) original UWA Narrative Corpus. The analysis commends two hypotheses for further testing: (1) speakers who use the NarrPP in storytelling are engaged in non-professional occupations (the NarrPP is negligible in the professional cohort) and (2) the NarrPP is more heavily used by males (85% [35/41]). These trends suggest that the AusE NarrPP lies at the heart of non-professional males’ vernacular and is likely to be a stable sociolinguistic variable holding covert prestige.

Keywords: Australian English, Narrative Present Perfect, Labovian narrative, sociolinguistic constraints.
1. Introduction

The English Present Perfect (henceforth PP) is not expected to occur in narratives since “the perfect is not a narrative tense” (Lindstedt 2000: 366) and the form has not been described as undergoing anterior-to-perfective grammaticalisation (Bybee, Perkins & Pagliuca 1994).¹ For these reasons, the (Standard) English PP does not generally collocate with definite past temporal adverbials and is not used to express temporal progression in narrative discourse. However, Engel and Ritz (2000), Ritz and Engel (2008) and Ritz (2007; 2010) report innovative uses of the PP in Australian English (AusE) radio chat-show programs, police and news media reports. They notably document the PP in radio narratives where the form is used to advance narrative progression. The Narrative Present Perfect (abbreviated NarrPP) is illustrated in (1) from Richard’s (under construction) UWA Narrative Corpus.²

(1) So I was standing on the corner next to where she hit him. He’s gone up on two wheels, gone one-eighty sort of heading back the direction he’s come, still on two wheels. The cars behind him have slammed on their brakes, just as he’s start-- completed the full three-sixty on two wheels. And he just landed. (Male, 36, chef)

The phenomenon has been analysed in terms of its semantics and pragmatics (e.g. Ritz & Engel 2008) but, to date, the sociolinguistic constraints operating on the NarrPP remain to be explored. This paper offers some preliminary sociolinguistic insights into the narrative uses of the have + past participle construction in AusE. The analysis rests on 24 narratives produced by four female and four male West Australians, born and raised in Perth, of professional and non-professional

¹ Cross-linguistic categories appear in small caps (e.g. “present perfect”) while language-specific tenses are written with initial caps (e.g. “Australian English Narrative Present Perfect”).

² In line with Walker (2011), the term “Narrative Present Perfect” (abbreviated NarrPP) refers to uses of the PP which occur in narratives and encode foregrounded story events which in turn advance temporal progression.
occupations and aged 36 to 56 years old. I identify the complicating action clauses which form the skeleton of the narratives, extract and code for tense the verbs heading those clauses and analyse tense variation between the forms used, focusing on the NarrPP. Preliminary findings indicate that the NarrPP is used mainly by non-professional males.

The paper is organised as follows. In Section 2 I introduce the English PP. I review the constraints traditionally associated with the use of the form in (Standard) English and present innovative uses of the English PP which have been reported in the literature. In Section 3 I introduce the narrative as defined by Labov and Waletzky (1967) and discuss the features of performativity and tense switching. In Section 4 I present the data and methods used and, in Section 5, I report some preliminary findings. I discuss their implications in Section 6 and conclude in Section 7.

2. The perfect

2.1 The (Standard) English Present Perfect

The present perfect establishes a link between a past situation and a current state, but is not “about” that past situation (Comrie 1976: 52). Therefore, definite past temporal adverbials indicating date and/or time, sequential adverbs such as ‘then’, or relational adverbs such as ‘when’ and ‘while’ are not permitted with the Standard English PP. This is illustrated in (2) where the collocation of the definite past temporal adverbial ‘yesterday’ with the PP is infelicitous, though the sentence is intelligible (Smith 1997: 108).

---

3 Professional occupations refer to occupations that require advanced education and are evaluated as more prestigious on Congalton’s (1963) scale of occupational prestige. These include lawyers, medical specialists and academics. Non-professional occupations do not require advanced education and are rated lower on the prestige scale. Examples of non-professional occupations include bricklayers, motor mechanics and salespersons.

4 There is ample discussion and controversy on the status of the English PP as a tense and/or an aspect in the literature (cf. Reichenbach 1947; Comrie 1976; Mittwoch 1988; Klein 1992; Michaelis 1994; Nishiyama & Koenig 2010 inter alia). The debate falls outside the scope of this study, but here I follow Smith (1997: 186) and analyse the PP as “a construction with special meanings of temporal location and aspect”.

~ 31 ~
Another constraint on the English PP is its purported inability to express temporal progression in narrative. As Hopper (1982) remarks, the perfective aspect – and not the perfect – is used when narrating sequences of discrete situations. In English, the Simple Past (SP) is the tense of narration (Givón 1982; Bybee et al. 1994: 55) since it is the form which yields perfective aspect reading when used with eventive predicates, as shown in (3).

(3) My brother used to sleepwalk when he was little. And the worst experiences- one time he came- woke up, c-- walked into my room, slapped me on the face, laughed and went, “Haha!” And then he just went back to his room and went back to sleep again. (Female, 36, stay-at-home-mother)

A cross-linguistic tendency has been reported for perfects to grammaticalise into past or perfective markers (Bybee et al. 1994).\(^5\) In French, for instance, the Passé Composé has grammaticalised to the extent that it now encompasses both perfect and perfective meanings. Therefore it can occur with definite past temporal adverbials (example 4) and be used in narrative sequences to express temporal progression (example 5).

(4) Hier matin je suis arrivé en retard au boulot. (Male, 53, pastry chef) 

Yesterday morning I arrived late at work.

(5) [...] Puis y’a un moment, euh la jeune maman elle s’est arrêtée dans l’escalier un peu brutalement, et elle, comme elle était derrière, elle l’a légèrement- euh si tu veux, voilà son coude a légèrement bousculé je crois le sac de Mme Delorme. […] (Female, 54, unemployed)

Then at some point er the young mother she stopped in the staircase quite abruptly, and she, because she was behind, she slightly- er if you like, you see her elbow slightly pushed I think Mrs Delorme’s handbag.

---

\(^5\) In some languages perfects grammaticalise into evidential markers (Bybee et al. 1994: 95), a path which the AusE PP might be following (cf. Ritz & Caudal 2008).

\(^6\) The data in examples (4) and (5) were obtained from recorded informal conversations conducted in 2015 with French native speakers.
In this regard, the documentation of the English PP to express temporal progression in narratives is not unexpected and suggests an evolution of the form in the varieties concerned.

2.2 Innovative uses of the English Present Perfect

Levey (2006) and Walker (2011) document innovative uses of the PP in British English (BrE), in, respectively, working-class preadolescent narratives, and in footballers’ language. Walker (2011: 78) argues that “some kind of personal or emotional involvement in the narration of the events being described” acts as a trigger in the use of the NarrPP. Pragmatically, the NarrPP seems to have a mirative function as it highlights unexpected events (Walker 2011: 77).


(6) Pernice, a legend of the ‘alternative country’ scene has toured the US […] and has also played in Australia in March 1999 to great response. (Engel & Ritz 2000: 137)

The AusE PP (NarrPP) also appears in “sequences indicating narrative progression” (Engel & Ritz 2000: 119), exemplified in (7) (Engel & Ritz 2000: 134).

(7) I’d done enough, and she said “Can you sign this?” and I said “Oh, okay, one final signing, I promise, and will you go away?” and she said “Yeah, yeah”. So I’ve got a texta, I’ve held her head straight and I’ve written on her forehead “Hi Mum, I’ve tried drugs for the first time”. (Triple J radio Sydney, 29 February 2000)
According to Ritz (2010: 38) there is an extension of the meaning of the PP inAusE and “the PP in spoken Australian English is clearly not a barrier to narration” (Ritz 2007: 142).

3. Narratives

3.1 The Labovian narrative

A narrative has been defined as “one method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which (it is inferred) actually occurred” (Labov 1972b: 359–360). Labov and Waletzky (1967: 28) formally define a narrative as “any sequence of clauses which contains at least one temporal juncture”. Temporal juncture separates two clauses which are temporally ordered with respect to each other and which cannot be interchanged without triggering a change in the semantic interpretation of the order of the events reported (Labov & Waletzky 1967: 25; Labov 1997: 399). Central to the organizational structure of the narrative is also “the most reportable event” (Labov 1997: 405), a climactic event which is “the semantic and structural pivot on which the narrative is organized” (Labov 1997: 414).

Labov and Waletzky (1967) partition the structure of oral narratives of personal experience into five basic sections:7

- Orientation
- Complicating action
- Evaluation
- Resolution
- Coda

---

7 The narrator may also begin his/her story with an abstract, viz. a short summary of the narrative, similar to a title, which announces what the narrative is going to be about.
Crucially, complicating action clauses participate in temporal junctures, i.e. they correspond to the core series of events that are sequential, causally related, and advance the story. They lead to the “most reportable event” which ends the complicating action section. The set of complicating action clauses following the most reportable event constitutes the resolution of the narrative (Labov 1997: 414).

3.2 **Performativity and tense switching in narratives**

A specific type of narrative has been distinguished in the literature and identified as “performed” (Wolfson 1978), “vivid” (Ritz & Engel 2008) or “reportive” (Maranhão 1984). A performed narrative features a number of performance devices, namely direct speech, asides, repetition, expressive sounds, sound effects, and motions and gestures (i.e. kinesic markers) (Wolfson 1978: 216). Only some of these features need to occur in a narrative for it to be deemed “performed”. What sets aside “performed” narratives is the enactment of the past events reported as opposed to a mere oral account of those events. The best narrators are endowed with “histrionic gifts”, adopting the tone of voice, facial expressions and body posture of the characters in their story (Maranhão 1984: 253).

A story is more likely to exhibit tense switching when it is fully performed. In her work on American English (AmE) storytelling, Wolfson (1978; 1979; 1982) finds that the Historical Present (HP) only occurs in narratives where the speaker “breaks through” into performance (see Hymes 1975). To observe the NarrPP, one therefore needs performed narratives which are more likely to exhibit tense switching.

The use of tenses in narrative discourse is divergent from the use of tenses in everyday conversation because the relation between time and tense is different (Fleischman 1990: 3). The Simple Past (SP) is the default tense to report past events. However, the HP is reported in narratives (Schiffrin 1981; Wolfson 1982), and so is the PP (Section 2.2), despite the general consensus that the form does not occur in this discourse genre since it is not suited for narrative progression (Labov & Waletzky 1967: 29). Tense switching in narrative is used to foreground unknown and unexpected events; it is also used to highlight evaluated events.
So tense alternation marks off different events within the story, giving it structure. It also works as an internal evaluation device (Schiffrin 1981; Silva-Corvalán 1983). Schiffrin (1981: 51) observes that 30% [381/1288] of the verbs heading complication action clauses are in the HP in contrast with only 3% [9/268] of the verbs heading orientation clauses. The HP occurs more frequently in the climax and build-up to climax of the narrative (Schiffrin 1981: 60).

4. Data and methodology

4.1 Corpus

In this exploratory study I analyse a sub-sample of narratives from Richard’s (under construction) UWA Narrative Corpus. The corpus consists of narratives obtained in semi-naturalistic interaction with the researcher in 2014 and 2015. The Narrative Corpus is built on the premise that the AusE NarrPP occurs in performed Labovian narratives. As mentioned earlier, Ritz and Engel (2008) find such innovative uses in radio narratives delivered with almost no interruption. The situation provides a monologic type of discourse rather than a conversational/interactive one (Romano, Porto & Molina 2013: 73–74) and is reproduced as part of the data collection process. Participants are engaged in storytelling via a list of prompts based on Labov’s (1984: 34–36) conversational modules and a selection of radio questions.

The speaker sample consists of four males and four females, aged 36 to 56 years old, born and raised in Perth, Western Australia. Four speakers (two males and two females) are professionals with university degrees: a clinical psychologist, an accountant, a sustainability officer and a registered nurse. The remaining four speakers are non-professionals: a process operator, a chef, a medical receptionist and a remedial massage and stretch therapist. Occupation is used as a proxy for social class: “if social class is determined by a combination of features, the single indicator that accounts for by far the greatest portion of the variance is occupation” (Ash 2002: 419).
4.2 Data extraction and coding

The variable context or envelope of variation (Poplack & Tagliamonte 1989: 60; Tagliamonte 2012: 10) in this study comprises all complicating action clauses in performed Labovian narratives. The linguistic variable under scrutiny is narrative tense.

Three performed (Labovian) narratives are extracted per speaker (see Section 3) (24 narratives in total); all the extracted narratives contain a well-developed complication. Minimal narratives (i.e. narratives made up of only two complicating action clauses) are excluded on the basis that their minimal length might preclude tense switching. The coding protocol used is in line with Levey (2006) and Rodríguez Louro and Ritz (2014).

Following the principle of accountability (Labov 1972a: 72), all tokens within the relevant variable system are considered. All complicating action clauses are extracted and the verbs heading those clauses are coded for tense. Verbs heading subordinate clauses are not considered as they introduce background information. In (8) the non-standard uses of the NarrPP (underlined for illustration purposes) are not counted; only the verb (‘ve done) in the main clause is.

(8) […] and as I’ve come back out the door I’ve done this Oscar-winning fake, “Oh, I’ve tripped” as I’ve opened the thing. (Male, 36, chef)

Ambiguous tense forms, as shown in (9), are likewise excluded from analysis.

(9) And I’ve grabbed what I thought was the right one, lifted it up enough to see it was the right one. (Male, 46, process operator)

Deciding whether one is in the presence of an elided auxiliary have or a Simple Past (SP) form is problematic unless the past participle in question is irregular, as

---

8 I focus on complicating action clauses since they have been reported as the prime location for the NarrPP in previous research (cf. Engel & Ritz 2000: 133–134).

9 The focus of this paper is on temporal progression rather than simultaneity (or elaboration). Subordinate clauses are excluded following Labov and Waletzky’s (1967: 21) and Labov’s (1997) contention that subordinate clauses do not enter into temporal juncture, i.e. are not relevant to the temporal sequence of narrative (though see Couper Kuhlen 1989 on ‘narrative temporal clauses’).
is the case with *gone* in (10). One cannot assume the elision of the *have* auxiliary since speakers may alternate between the NarrPP and the SP, as shown in (10) (see Ritz & Richard in preparation).

(10)  [...] a huntsman spider [...] **has come** sort-of off the hand of this tool, and **ran up** 'cross my hand and **gone** under the sleeve of my um overalls.  
(Male, 46, process operator)

Also excluded from analysis are zero (or null) quotatives (Mathis & Yule 1994), as they cannot be coded for tense. This is illustrated in (11).

(11)  And I said, “Mel, shush. What are you laughing at? That’s rude like.”  
**ZERO** “Mum, they won’t even notice.  *<LAUGHING>* You can turn around and look.”  (Female, 53, nurse)

These protocols resulted in the extraction of 335 narrative clauses, each headed by a finite lexical verb.

### 5. Preliminary findings

The overall tense distribution in the sample is as follows. The SP is used in 68% [227/335] of the complicating action clauses, remaining the narrative tense par excellence. The HP features in 20% [67/335] of the complicating action clauses while the NarrPP follows suit, occurring at 12% [41/335].

How do these findings compare against existing research? The preponderance of the SP is in line with previous analyses on American, British and Australian English narratives (Schiffrin 1981; Levey 2006; Rodríguez Louro & Ritz 2014). The proportion of HP and NarrPP varies. A comparison of tense variation in narratives across three English varieties and based on different samples is plotted in Figure 1.10

---

10 The results presented in this figure stem from, in order of appearance (from left to right), this study, Schiffrin (1981) for American English, Levey (2006) for British English, and Rodríguez Louro and Ritz (2014) for Australian English.
On a sample of 73 American English (AmE) narratives, Schiffrin (1981: 51) identifies 30% [381/1288] of HP clauses, with 70% of clauses in the SP [907/1288]. She does not report any tokens of the NarrPP. Levey (2006), analysing tense variation in 56 preadolescent BrE narratives (ages 7-11), finds 60% of clauses in the SP [340/571], 32% in the HP [181/571] and 9% in the NarrPP [50/571] (Levey 2006: 140). For AusE, Rodríguez Louro and Ritz (2014) analyse 100 narratives of personal experience produced by 12 to 62 year-old speakers. The SP dominates complicating action clauses at 87% [571/654], the HP features at 12% [76/654] while the NarrPP is virtually absent at 1% [9/654] (Rodríguez Louro & Ritz 2014: 556).¹¹

The overall distribution of narrative tenses in the present study conceals important disparities in the sample. Table 1 displays the proportion of the

---

¹¹ The Ns from Rodríguez Louro and Ritz (2014) actually add up to 656 but I present the results here as they are reported in their paper.
different tense forms among professional and non-professional speakers separately.12

<table>
<thead>
<tr>
<th></th>
<th>Professional</th>
<th>Non-professional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Simple Past</td>
<td>152</td>
<td>94</td>
</tr>
<tr>
<td>Historical Present</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Narrative Present Perfect</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>161</td>
<td>100</td>
</tr>
</tbody>
</table>

\[ X^2 = 103; \ p = .000 \]

Table 1. Distribution of tenses heading complicating action clauses according to occupation

Professional speakers evince a system relatively invariant. The SP dominates at an overwhelming 94% [152/161]. This is comparable to Rodríguez Louro and Ritz’s (2014: 556) results of 87% SP across age cohorts. Rodríguez Louro and Ritz (2014) are unconcerned with occupation as a social variable, but their overall corpus is likely to be representative of speakers from higher socio-economic backgrounds (Rodríguez Louro & Ritz 2014: 553). Table 1 also shows that the HP only represents 5% [9/161] of the complicating action clauses of professional speakers and that the NarrPP does not feature at all. A chi-squared test for independence shows that there exists a statistically significant relationship between tense and occupation \( (X^2 = 103, \ p < 0.001) \).

In contrast to professional speakers, non-professional speakers show a system rich in tense variation. Antithetically to the system evinced by the professional speakers, the SP no longer dominates, occurring in the non-professional cohort at a modest 43% [75/174], less than half the times this form is attested in the professional corpus. The HP features at 33% [58/174], and the NarrPP occurs in 24% [41/174] of the complicating action clauses analysed.

\[ ^{12} \mathrm{Note\ that\ the\ percentages\ for\ professional\ speakers\ only\ add\ up\ to\ 99\%\ because\ of\ rounding.} \]
The differences in storytelling between professionals and non-professionals are exemplified in (12) and (13). The verbs heading complicating action clauses appear in bold.

(12) And um we checked into the hotel. […] And we got to our room and there were… twin- Yeah, two twins beds. And we thought, “That’s not right. We- we booked with a travel agent.” So we went down to the- to the- the reception desk […] And my wife’s going, “But it’s our honeymoon, we- we wanted a double bed when- why have we got single beds?” And they jabbered away in Indonesian […]. And then they said, “Oh we’ll- we’ll fix. We’ll fix. Go back to your room, we’ll fix.” […] And er next thing there’s a knock on the door and um they came in and they- […] So they took the unit out and they pushed the two beds together. Then they put another double mattress on top of- of the- the single beds there […] and then because they didn’t have double sheets so they folded the sheets somehow to make the sheets fit. (Male, 56, accountant)

In (12), the narrator builds the events of the story using the SP. There is only one occurrence of the HP in the progressive (‘s going) and no use of the NarrPP.\(^\text{13}\) The use of tenses ‘conforms’ to the temporality of events – past events are reported using the SP, the quintessential narrative tense. The narrator performs his story reporting direct speech and thought (they said, “…”; we thought, “…”) but his tense usage remains mostly invariable.

(13) I remember I was helping my dad my um- fix his tractor once […] And um he’s asked for a- some bloody tool. And I’ve put my hand in the toolbox trying to get up enough to see where it is, reaching around like that. And I’ve grabbed what I thought was the right one, lifted it up enough to see it was the right one. And […] a huntsman spider- […] has come sort of off the end of this tool, and ran up ’cross my hand and gone under the sleeve of my um overalls, […] So immediately I’ve thrown the spanner, er gone to get out, smashed my head an absolute beauty […] Dad’s yelling at me, “What’s going on? What’s going on?” And I’m not even answering him because I’m- I’m already- I’m already running and peeling off these um <LAUGHTER> these overalls at the same time,

\(^{13}\) I do not discuss the role of the progressive aspect in this paper. The progressive’s primary function in English is to signal an on-going process (Dahl 1985: 28). Comrie (1976: 25) views the progressive aspect as a subcategory of the imperfective. Therefore the progressive is usually restricted to clauses referring to backgrounded events. However, the progressive has been reported in sequences expressing temporal progression (see Couper-Kuhlen 1995; Mesthrie 2013).
doing the whole <MIMICKING HIS AGITATION IN REMOVING THE OVERALLS> girl thing you-know <SCREAMING> “Get off me!” (Male, 46, process operator)

In (13) the narrator resorts to the NarrPP rather than the SP to advance the story. There is only one token of the SP (ran up) and one token of the HP in the progressive (‘s yelling). There are also two ambiguous forms (lifted and smashed) which, as explained earlier, were excluded from the quantitative analysis. The narrator uses mimesis in his story, as indicated by the elaboration “reaching around like that” and the information presented in angled brackets (<MIMICKING HIS AGITATION IN REMOVING THE OVERALLS>). He also resorts to direct speech, reporting his father’s words (Dad’s yelling at me, “What’s going on? What’s going on?”).

My findings are right in line with Levey’s (2006) observation that 9% of the complicating action clauses of working-class preadolescent narratives featured the NarrPP. Levey’s study focuses exclusively on working-class storytelling; the findings in this study, however, are reflective of a deep schism in how professionals and non-professionals deploy tense switching in narrative. The findings show that, as with Levey’s (2006) finding for BrE, AusE narrative tense switching is constrained by speakers’ occupational background.

Tense distribution in the sample is also plotted against the sex of speakers. This is represented in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Simple Past</td>
<td>84</td>
<td>65</td>
</tr>
<tr>
<td>Historical Present</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Narrative Present Perfect</td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>TOTAL</td>
<td>130</td>
<td>100</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 51.9; p = .000 \]

Table 2. Distribution of tenses heading complicating action clauses according to sex
The proportion of SP tokens is similar across sex (65% [84/130] for male speakers; 70% [143/205] for female speakers). The second most commonly used tense form for male speakers is the NarrPP at 27% [35/130] followed by the HP at 8% [11/130]. Conversely, the second most commonly used tense form for female speakers is the HP at 27% [56/205] followed by the NarrPP at a low 3% [6/205]. Female speakers barely use the NarrPP and prefer the HP. On the other hand, male speakers’ narratives feature the HP at a modest 8%. The differences between the expected and observed distribution of tenses according to sex are significant ($\chi^2 = 51.9$, $p<0.001$). The NarrPP represents more than a fourth of the complicating action clauses of male speakers, but the distribution of tenses according to occupation in Table 1 revealed that the NarrPP was only niched in the non-professional cohort: all tokens of the NarrPP stem from the non-professional cohort and 85% [35/41] of those tokens belong to male speakers.

6. Discussion

Occupation and sex have emerged from the analysis presented above as two social variables significantly contributing to tense switching and the use of the NarrPP in particular. The narratives of non-professionals display clear variability with the SP occurring in less than half of the complicating action clauses (43%) while the HP and the NarrPP respectively feature at a rate of 33% and 24%. These findings are unique and only possible due to the nature of the sample, consisting of performed Labovian narratives produced by females and males of various occupational backgrounds. Previous studies of AusE, namely Rodríguez Louro and Ritz (2014: 557), have documented high usage rates for the SP (87%) and lower usage rates for the HP (12%) and the NarrPP (1%) and these results are likely to stem from the relative homogeneity of their corpus: the data largely stem from students in conversation with friends and family and are generally representative of educated, professional speakers.14 The importance of speakers’ socio-economic status is reinforced by Levey’s (2006) findings on a dataset of BrE working-class preadolescent narratives. The variability displayed and the non-

14 The University of Western Australia is located in the Western suburbs of Perth (an affluent neighbourhood according to the Australian Bureau of Statistics) and attracts students from the area.
negligible occurrence of the NarrPP (60% SP, 32% HP and 9% NarrPP) is in line with the results of the non-professional cohort in this study (68% SP, 20% HP and 12% NarrPP). Speaker occupation thus emerges as a potential social constraint on tense variation and the use of the NarrPP.

The usage patterns observed across speaker sex show that men are the main users of the NarrPP. This result also resonates with Levey’s (2006) BrE data where the NarrPP represents only 6% [20/312] of the complicating action clauses of female preadolescents but 12% [30/259] of the complicating action clauses of male preadolescents. Men are using the NarrPP, a non-standard variable, more often than women.\textsuperscript{15} According to Labov’s Principle I, men use more non-standard forms than women regarding stable sociolinguistic variables (Labov 1990: 210). On the other hand, “women show a lower rate of stigmatized variants” for such variables (Labov 2001: 266). Female speakers are more likely to use prestigious/standard linguistic forms as a result of the social pressure to conform to norms and as a way to secure their social status (Trudgill 1972: 182–183; Romaine 1978: 156). The distribution across sex of the non-standard NarrPP suggests that we are in the presence of a stable sociolinguistic variable.

The NarrPP being a non-standard variable used almost exclusively by non-professional males, the question arises whether the form may act as a covert prestige marker. Trudgill (1972: 179) argues that “working-class speech has favourable connotations for male speakers”. Unsolicited metalinguistic commentary by a 31-year-old female teacher in the UWA Narrative Corpus describes the NarrPP as the domain of “blokey males”. Ellis (2012) finds that attitudes (from university-educated respondents) towards the innovative uses of the PP in AusE narratives and police media reports are generally negative. NarrPP users are primarily associated with the blue collar/“not a professional category”, and with lower socio-economic groups (Ellis 2012: 42, 46). The educational level

\textsuperscript{15} I consider the use of the HP in narratives as standard; the use of the NarrPP as non-standard. The use of the HP in narratives to refer to past events is stylistically marked but is a well-accepted usage (see Dahl 1985: 14; Kearns 2011: 180). While prescriptive grammars of English list the genre-specific use of the present tense in narration, they never mention a similar use for the present perfect.
of NarrPP users is evaluated negatively (Ellis 2012: 71). However, a linguistic feature may be negatively evaluated by the speech community, yet represent a source of covert prestige and lexical solidarity for a certain social group (Guy, Hovarth, Vonwiller, Daisley & Rogers 1986: 38).

While previous research (Ritz 2010: 3412; 2012: 899) proposed that the NarrPP might represent a change in progress in AusE, the results of the current analysis show that – synchronically – the NarrPP is the preserve of the non-professional male. The restriction of the NarrPP to the non-professional cohort rather points towards a stable sociolinguistic variable. An apparent time study is however needed to confirm this hypothesis.

Levey (2006: 148) suggests that the NarrPP is a pragmatic phenomenon rather than a grammaticalised substitute for the SP. The use of the NarrPP is similar to that of the HP in that it demarcates events in the narrative (Schiffrin 1981; Fleischman 1990). In the findings for the non-professional cohort, the proportion of HP and NarrPP usage between male and female speakers is reversed: for males the proportion of NarrPP is 49% [35/71], that of HP is 14% [10/71]; for females the proportion of NarrPP is only 6% [6/103] but that of HP is 47% [48/103]. This suggests that the NarrPP and the HP fulfil similar functions in performed narratives.

7. Conclusion
Although the English PP has not been reported to grammaticalise into a past or perfective marker (Bybee et al. 1994), its deployment in AusE narratives, specifically in complicating action clauses, has been well documented. Yet, the sociolinguistic constraints on its use remained to be established.

An exploratory sociolinguistic analysis of a sample of 335 complicating action clauses stemming from 24 performed narratives indicates that the NarrPP is significantly used by male speakers of non-professional occupational backgrounds. Crucially, the NarrPP is virtually absent from the narratives of the professional cohort and the latter display a lack of variability in the choice of
tenses, overwhelmingly resorting to the SP (94%). By contrast, the narratives of the non-professional speakers are replete with variation, the HP and NarrPP respectively featuring at 33% and 24%.

Previous research suggested that the use of the NarrPP might reflect a change in progress (Ritz 2010; 2012). Though in the absence of cross-generational data one cannot answer this claim, the foregoing analysis has revealed that the use of the NarrPP is restricted to the non-professional cohort. Moreover, it is largely preferred by males (85% among non-professionals). This hints at the possibility that the NarrPP represents a stable sociolinguistic variable with covert prestige for its users, namely non-professional males.

Multivariate statistical analysis on a larger dataset is the next step to establish the ranking and strength of the social and linguistic constraints on the NarrPP in AusE performed narratives as well as to explore lexical effects on the use of the SP, the HP and the NarrPP (Richard in preparation; Rodríguez Louro & Richard in preparation; Richard & Rodríguez Louro under review). The results of a statistical analysis spanning several generations of speakers will also allow us to ascertain the status of the NarrPP as a stable sociolinguistic variable as opposed to a change in progress.

Acknowledgements

I gratefully acknowledge the support of the Australian government and the University of Western Australia for their funding of my Ph.D. research through an International Postgraduate Research Scholarship (IPRS), an Australian Postgraduate Award (APA) and a UWA Top-up Scholarship. I also wish to thank an anonymous reviewer and the Editors for their comments and suggestions on this paper. Any errors remain my own responsibility.

References


Congalton, A A 1963 Occupational status in Australia Kensington, NSW: School of Sociology, University of New South Wales.


Cox, R 2005 Preterite uses of the present perfect in New Zealand English narratives: A case study The University of Canterbury: M.A. thesis.


Ellis, K 2012 ‘That’s the way they’re used to speaking’: Attitudes towards innovative uses of the present perfect in Australian narratives and police media reports The University of Western Australia: Honours thesis.


Hopper, P 1982 Tense-aspect: Between semantics and pragmatics Amsterdam/Philadelphia: John Benjamins.


Richard, S in preparation *Sociolinguistic constraints on the Narrative Present Perfect in contemporary Australian English* The University of Western Australia: Ph.D. thesis.

Richard, S under construction ‘UWA Narrative Corpus’ Discipline of Linguistics, The University of Western Australia.

Richard, S & C Rodríguez Louro under review ‘Narrative-embedded variation and change: The sociolinguistics of the Australian English Narrative Present Perfect’ In V Werner, E Scoane & C Suárez-Gómez (eds) *Re-assessing the present perfect* Berlin: Mouton de Gruyter.


Ritz, M-E & S Richard in preparation ‘The functions of the auxiliary have in Australian English vivid narratives’ Paper to be presented at the *2008 Conference of the Australian Linguistic Society University of Sydney*, July 2008.

Ritz, M-E & D Engel 2008 ‘Vivid narrative use and the meaning of the present perfect in spoken Australian English’ *Linguistics* 46(1): 131–160.


Switch-reference and the subject in New Guinea:  
the case of Yagaria

Glenn Windschuttel  
University of Newcastle  
g.windschuttel@gmail.com

Abstract. Switch-reference in New Guinea has been claimed to follow topics or some pragmatically prominent position (Donohue 2005; Roberts 1997). That the subject (as determined, for example, by verb agreement) is not always followed was noticed early on (Longacre 1972). This has a particular form sometimes called ‘clause skipping’ (Roberts 1987:299). Examples in the texts in the Hua Grammar (Haiman 1980) show this exists in Yagaria, a Trans-New-Guinea language (of which Hua is a dialect).

In the object experiencer construction the ‘subject’ according to verb agreement is also ignored in clause-chains in preference to the experiencer. In many languages this fits into the ‘clause skipping’ pattern and this suffices as an explanation (Donohue 2005; Roberts 1997). However, in Yagaria/Hua this is not the case. This switch-reference, instead, indicates that the experiencer is the subject, albeit, one bearing quirky case. Furthermore, there is some evidence that their object-like morphology derives from a time when they were true objects (supporting a claim about the acquisition of subjecthood made by Cole et al 1980).

Keywords: Papuan, experiencer, switch-reference, subject
1. Introduction

In the Papuan languages the semantic notion of actor maps almost directly onto subject. The actor and subject are not separated by voice alternations such as the passive which are basically unknown among the Papuan languages (Foley 1986, 2000). Nor in many languages are there functional relations like control or raising that could necessitate a subject removed from semantics.

The impersonals (also known as object experiencer constructions; Evans 2004) present difficulties for the interpretation of subjecthood in Papuan languages. This is because the proto-agent properties (Dowty 1991) are shared by both participants. The stimulus has no proto-patient properties and is depicted as the cause. The affectee/experiencer is sentient/perceiver while also having the proto-patient properties of being casually affected and undergoing a change of state.

For this reason grammatical properties usually associated with the subject are also shared by both participants (which of these properties are crucial to the definition of the subject varies from language to language). Example (1) from Kalam (Madang) illustrates this. In some sense the excrement that needs to be voided causes the feeling/need that affects the experiencer; however, it is the experiencer who is the perceiver and is sentient. This is reflected in the grammar. The experiencer regularly appears first in the clause as is normal for subjects but uses accusative pronouns normally confined to objects. Similarly despite being positioned second in the clause the stimulus is the only argument possibly encoded on the verb. Other characteristics of impersonals will be discussed in section 2.

(1) KALAM yp sb yow-p
1SG.ACC excrement fall-PFV.3SGA
I need to defecate/I feel like defecating (Pawley et al. 2000:168)

Switch-reference is one notion that is typically stated in terms of the subject: ‘same subject’, ‘different subject’ and ‘anticipatory subject’ (section 3). The
system, however, is rarely this straightforward: switch-reference may ‘skip’ pragmatically unimportant actors and mark as same subject clauses that do not share participants. A fuller description of this phenomenon is given in section 3.1. It is the interplay between switch-reference and impersonals that is particularly interesting. This is a point of variation among the Papuan languages: the experiencer is often given extra importance but how varies. This will be briefly surveyed in section 3.2. This has previously been seen as unimportant in defining grammatical relations in impersonals (see section 4); however, this will be challenged by data from Yagaria.

Switch-reference will provide evidence for a subject distinct from the actor in Yagaria, a Trans-New-Guinea (TNG) language of Papua New Guinea (PNG) that will be introduced in section 5 (Haiman 1980 wrote a famous grammar of one of its dialects, Hua). This non-actor subject is only clear in the impersonals described in section 6. This is established through switch-reference, the non-canonical forms of which are examined in section 7, and fully justified in section 8. The same construction can have its grammatical relations differently organised with the experiencer as object befitting its coding (section 9). Moreover, there is some support for this interpretation being historically the older of the two (section 10).

2. Impersonals

Impersonal constructions are a common method for expressing uncontrolled experience. These expressions have two general features:

- An experiencer which resembles objects.
- Subject agreement (if present) is third singular.

There are two further properties that are routinely found in impersonals:

- A nominal that is somewhat idiomatically fixed in the construction.
- The experiencer (when an NP) appears in the position typical of subject.

Example (2) from Yagaria illustrates some of these properties. Objects in Yagaria are rendered by prefixes: this is the case for the first person experiencer below as
The third singular subject agreement is zero in this case. There is also an idiomatically fixed stimulus nominal, *agaso* ‘itchiness’.

(2) MOVE¹

| agaso | no-d-ei-Ø-e |
| itchiness | PROG-1SGU-hit-3SGA-IND |

I am itchy lit. itchiness hits me (agaso ei- entry in Renck 1977)

These constructions exist in all parts of New Guinea, even outside of the TNG languages. They are found in the extreme northwest on the Bird’s Head, for example, Sougb in (3); along the spine of the Highlands (the heartland of the TNG languages), for example, Yagaria above in (2); and even down south among the Yam languages such as Nen in (3). The non-TNG languages of the North are not excluded, for example, Kombio of the SVO Torricelli languages in (5). All these languages can be found on the map below, Figure 1: Sougb as a blue dot; Yagaria inside the green ellipse (see Figure 2); Nen as the yellow dot towards the bottom; and Kombio, the red dot in the middle of the north coast. Similar expressions are also widespread across the Torres Strait in northern Australia in both the non-Pama-Nyungan languages (Evans 2004; Walsh 1987) and the Pama-Nyungan languages (Verstraete 2011).

(3) SOUGB

| dan | ar-eb-ed |
| 1SG | thing-do-1SGU |

I am sick (Reesink 2005:192)

(4) NEN

| ynd | ak-ām | w-ram-t-e |
| 1ABS | thirst-ERG | 1SGU-make-NDU:IPFV-3SGA |

I am thirsty (Evans 2011:11)

(5) KOMBIO

| mantiek, | nilmp | a-r | apm |
| mother | stomach | REALIS-hit | 1SG |

*yat-ip!*

very-LIM

mother, I’m very hungry! (lit. ‘My stomach is hitting me very much!’; Henry nd:22)

¹ The Yagaria data are specified for which dialect they originate from, Move or Hua.
We can see from these examples some of the ways the experiencer is treated similar to objects. In most cases presented here undergoer verb morphology is used: (2), (3), and (4). The focus of this article will be on Yagaria that also displays this behaviour. However, this is not the only means: in (5) it is the postverbal word order that groups the experiencer with objects and in (1) we saw it was the pronominal paradigm used.

In all these examples there is also a stimulus nominal, a possible subject; in (4) this is clearly the case, the stimulus nominal receiving ergative case while in Yagaria this will turn out not to be the case (see section 6).

In the Sough, (3), and Nen, (4), examples the experiencer is given greater prominence by word order. Both are SOV languages but in both examples the experiencer, otherwise treated like an object, precedes the stimulus nominal.

This quick overview reveals some of the diversity that exists in this group of constructions. Diversity of a different sort will be discussed in the following sections; how they behave in relation to switch-reference in interclausal syntax and the implications that has for grammatical relations in impersonal clauses.
3. Switch-reference

Clause-chaining is one of the typical ‘Papuan’ traits (Foley 1986, 2000, Reesink 1983). Many languages have dedicated morphology for the medial verbs (the non-final verbs in the chain). These verbs are often somewhat simplified and depend on the final verb for full expression of tense, aspect and mood features.

In many languages the medial verb inflects for the continued presence of a particular participant in some grammatical or pragmatic role across clause boundaries, a system of switch-reference. Most commonly switch-reference is said to mark whether the subject of a given clause and the following are the same, using same-subject marking (SS) or whether they are not, using different-subject marking (DS) (Haiman & Munro 1983; cf: Roberts 1997). This is clear from the following examples in (6) from Usan, a Madang language of PNG. In (6a) the subject of both clauses is the same and we see the same-subject morpheme, -ab, suffixed to the medial verb while in (6b) the subject changes between clauses and the different-subject morpheme, -ine, is used.

(6) USAN

\[a. \text{ye nam su-ab isomei}\]
\[1SG \text{ tree cut-SS go.down.1SGA}\]
\[I \text{ cut the tree and (I) went down}\]

\[b. \text{ye nam su-ine isorei}\]
\[1SG \text{ tree cut-DS go.down.3SGA}\]
\[I \text{ cut the tree (and it went) down (Haiman & Munro 1983:viii)}\]

The switch-reference systems of the Papuan languages vary greatly as established by Roberts (1997), a survey following similar studies of Australian (Austin 1981) and Amerindian languages (Jacobsen 1983).

This variation is on a number of parameters. The first parameter to be considered is the place of the system itself in the grammar. On one end of the spectrum are languages like Usan where the marking is compulsory and the meaning inherent to the marking. These can be contrasted with less rigid systems like Mian (Fedden 2012) and Dani (Bromley 1981) which are largely based on implicatures. This paper will be devoted to those languages where the system is highly grammaticalised.
Even among these languages there is a great deal of diversity. Morphologically they can be as simple as Usan above indicating basically only the identity/non-identity of subjects or as complex as the systems of the Gorokan family. In this family the SS marker, not only marks the subject as the same, but also inflects for that subject’s person and number. DS markers anticipate the subject of the following clause encoding its person and number features as well as indicating that it is different to the preceding subject.

Yagaria is a Gorokan language. Examples of both situations, DS and SS, are below, (7a) and (7b) respectively. The DS marker in (7a), -agada, indicates not only that the subject is not shared by the two clauses but also that the following subject is first singular. The morpheme -na in (7b) indicates that the subject is the same throughout and also that it is third singular. The form of the markers themselves will be examined in greater depth in section 7.

\[
(7) \quad \text{MOVE} \quad a. \ ba \quad d-ami-d-an-agada \\
\text{sweet.potato} \quad 1\text{SGU-give-PST-2SGA-DS-1SG} \\
no-d-o-e \\
\text{PROG-eat-1SGA-IND} \\
you gave me a sweet potato I am eating it (Renck 1975:115) \\
b. \ dope'na \quad no-do-na \\
\text{food} \quad \text{PROG-eat-3SG.SS} \\
ge \quad hu-d-i-e \\
\text{word} \quad \text{do-PST-3SGA-IND} \\
while he was eating he spoke (Renck 1975:108)
\]

In clause-chains the subject is tracked between clauses by morphemes that vary according to whether it is the same or not. The details of the form of this morpheme and what other features it encodes vary from language to language.

3.1 ‘Clause skipping’

The summary above is somewhat oversimplified. Not only can the morphology be complex but the participant that is tracked can vary within a language. The subject can be partially ignored in preference to some form of topic.

A clause chain may use SS morphology despite the presence of a different, but pragmatically irrelevant, subject following ‘clause skipping’; Roberts 1987:299; see
also Longacre 1972:12-14; Roberts 1997:174-6; Donohue 2005; Donohue 2008; Wade 1997). This is always of one particular form given in (8) below. That is, there is a sequence of three clauses, the first and third of which share the same subject, while the second clause has a distinct subject (each clause is bracketed and numbered with a superscript). Following the general principles of switch-reference the expectation would be that both verbs would be DS marked; however, in clause skipping the first verb is instead marked SS.

(8) \[ [\text{sbji } \text{V-SS}]^1 \ [\text{sbji } \text{V-DS}]^2 \ [\text{sbji } \ldots]^3 \]

The Buin example in (9) exemplifies this. The subject of the first and third clause is first plural while ‘rice’ is the non-topical subject of the second clause. The same-subject marker precedes the clause with the different subject. This clause is then marked as having a different-subject from the following in which the original subject resumes its role.

(9) BUIN \[
\begin{align*}
[\text{egu } \text{iko } \text{raiti } \text{kuruin } \text{eeto-gi-mo}]^1 \\
\text{now that rice sow } \text{PLA.do-SS} \\
[\text{egu } \text{kinatuguraa-gu}]^2 \\
\text{now 3SGA.grow-DS} \\
[\text{aapotogigu}]^3 \\
\text{PLA.plant.3SGU} \\
\text{now we sow that rice and when it comes up we plant it}
\end{align*}
\]

(Longacre 1972:42-44)

It should be noted that even in clause skipping the subject is always of importance to the system. Although the change to the irrelevant subject is actively ignored by the first verb, the second switch-reference morpheme does make reference to the ‘irrelevant’ subject. It is marked DS because of the change from non-topical subject to the resumed topical subject of the third clause.

In clause skipping an otherwise ungrammatical sequence of switch-reference markers has taken on the role of backgrounding a non-topical subject to keep the flow of discourse.
3.2 Switch-reference and impersonals

Impersonals are often referenced in a particular way in clause-chains. According to Roberts (1997 and references therein) this varies from language to language. There are two major patterns: these are given below in (10) and (12).

The first links a preceding ‘personal’ subject with the impersonal experiencer. This same argument is ignored when compared with a personal subject following. Amele, Usan, Telefol, Yau (Roberts 1997:169), and Lani (Donohue 2005) display this pattern which is exemplified for Usan in (11). The worker, the subject of the first clause, experienced exhaustion in the second clause and stops work in the last clause. The subject of the first clause is marked as same subject when compared with the apparent object, the impersonal experiencer of the second clause. The last subject is marked as the different from the subject of the impersonal clause.

The other pattern has the DS and SS marking in the opposite order. Yagaria follows this pattern. This is exemplified in (13), which actually displays two clause chains: (13a) shows a personal clause followed by an impersonal clause while (13b) has the personal clause following the impersonal one. Only combining these two clause chains together do we get the pattern in (12) (thus in some sense its existence as one unified clause chain is hypothetical).

(10) [sbji V-ss][exp V-DS] [sbji ...]

(11) USAN

\begin{verbatim}
munon isig eng sarau äib eb-ef1
man old ART work big do-ss
migeri wär-a2
exhausting 3SGU.hit-DS.3SG
weg-ib-ä3 cease-FUT.SGA-3SGA
\end{verbatim}

the old man will work hard and become exhausted and he will stop (Reesink 1987:203-4)
(12) [sbj V-DS][exp V-SS][sbj ...]

(13) MOVE

a. [hoya ol-u-ga-ni] garden take-1SGA-DS-3SG
[da-bag-O-e] 1SGU-be.well-3SGA-IND
I am working and like it (hago- entry in Renck 1977)
b. [gei-da da-bago-da] sickness-1SG 1SGU-recover-1SGS
[hava’a bei-d-u-e] just.so live-PST-1SGA-IND
I recovered and lived just so (Renck 1975:147)

Thus according to switch-reference the experiencer is treated like no ordinary object in these languages. (In both examples there was an optional pattern where this need not be the case and both verbs are marked DS as if the impersonal experiencer is a regular object.) However, for the reasons given in the next section this has not been seen as significant for establishing its grammatical function.

4. Previous accounts of the grammatical relations in impersonals

Both Donohue (2005) and Roberts (1997) conclude that the experiencer in impersonals has the grammatical relation befitting its coding, object. Its position at the start of the clause and the unusual switch-reference is entirely due to the topicality of the experiencer.

In Lani, argues Donohue (2005), there exists a clause initial pragmatic position to which objects may be moved. Example (14a) is an instance of this pattern: the third singular pronoun at representing the object comes first followed the subject nggewo ‘dog’. In (14b) we see the same structure in an impersonal clause. The pronoun at representing the experiencer is again clause initial and is also of clear pragmatic significance. For this reason, the existence of such a pragmatic slot is sufficient to explain the position of the experiencer before the stimulus nominal.

(14) LANI

a. at nggewo nen kege
3SG dog ERG 3SGU.3SGA.saw
the dog saw him (Donohue 2005:182)
b. at andi ekerak
3SG sickness 3SGU.3SGA.did
he was sick (Donohue 2005:197)

The other anomalous behaviour relates to switch-reference. This was discussed in section 3.2. The similarity between ‘clause skipping’ and the first pattern of impersonal switch-reference (SS-DS) is undeniable. Schemata for the two are repeated below in (15) and (16) respectively. Examples of these two clause chain types can be seen in (9) and (11) in the preceding sections. Both involve the same marking; moreover, in both cases the subject is shared by the first and third clause while the middle clause has no new topical participant. For this reason Donohue (2005) and Roberts (1997) both propose that switch-reference in impersonals is an example of clause skipping and this suffices as an explanation of this behaviour.

(15) \[\text{[sbji } V\text{-SS]}^1 \text{ [sbji } V\text{-DS]}^2 \text{ [sbji } \ldots ]^3\]

(16) \[\text{[sbji } V\text{-SS]}^1 \text{ [expi } V\text{-DS]}^2 \text{ [sbji } \ldots ]^3\]

This is a coherent account of the impersonals of these languages. The extra syntactic prominence given to the object experiencer simply follows general techniques available for marking arguments as pragmatically prominent. The clause initial position is available to all arguments that are pragmatically significant and this option is taken by the experiencer in impersonal clauses. A particular form of non-canonical switch-reference, clause skipping, allows pragmatically insignificant subjects to be ignored in clause chains. The non-experiencer argument in impersonal clauses is ignored in this way. No extra explanation of impersonals is necessary.

However, this explanation does not cover languages, such as Yagaria, that display the DS-SS pattern of switch-reference for their impersonals. Donohue (2005) is silent on this point while Roberts (1997) sees a difference in the whole switch-reference system of these languages. He suggests that the explanation of impersonal switch-reference is the same in these languages. The only difference is that this pragmatic pattern of switch-reference works in the opposite way, i.e., (17). He did not know of the existence of clause skipping in any of these languages. However, for a similar unified theory that explains impersonals as part
of a general pattern of pragmatic switch-reference, clause skipping would need to follow this same DS-SS pattern, i.e., (18).

(17) [sbji V-DS]₁ [expi V-SS]² [sbji ...]³
(18) *[sbji V-DS]₁ [sbjj V-SS]² [sbji ...]³

We will see in later sections that this is not the case. In section 7 it will be evidenced that clause skipping does exist in at least one DS-SS language and instead it follows the same pattern common to SS-DS languages, i.e., (15). Hence, no such unified theory of non-canonical switch-reference is possible. The consequences of this will be examined further in the sections that follow. Yagaria is the DS-SS language from which this evidence comes.

5. Yagaria

Yagaria, as mentioned in section 3.2, is one language with DS-SS switch-reference for impersonals and, it will turn out, also has clause skipping. The language will be introduced in this section.

The language Yagaria is spoken in the heartland of the TNG family in the Eastern Highlands province of Papua New Guinea. More specifically it belongs to the Kainantu-Gorokan family (Lewis et al 2015). The location of the language is shaded in Figure 2. Figure 2 corresponds to the ellipse in Figure 1, giving the language’s location in the wider context of the whole island. The dialect cluster was spoken by approximately 21,100 speakers in 1982 (Lewis et al 2015). Substantial grammars exist for two Yagaria dialects: the largest dialect, Move (Renck 1975), and Hua (Haiman 1980). In addition there are extensive dictionaries for these two dialects (Renck 1977 and Haiman 1991, respectively). This study will focus on Move; Hua data will supplement gaps as needed.²

² There is only some dialectal variation. All examples will be given in original orthography (one apparent difference is artificial: Hua is written without predictable vowels while Move opts to write them in).
Yagaria is a typical Papuan language with SOV word order, extensive use of SVCs and clause-chaining. There is complex person morphology that can show agreement for both subject and object. The subject marker is suffixal, often fused with tense morphemes and the root in complex ways, while object agreement is by a simple prefix. This is illustrated by (19)

(19) MOVE \text{\texttt{pa-va’\textlangle no\textrangle k-ou-e}} \\
\text{\texttt{3PLU\texttt{-slaughter\texttt{-PROG}}\texttt{-1SGA\texttt{-IND}}} \\
I am slaughtering them (Renck 1975:195)

The morphophonology of the agreement suffix means that it is too complex for a reasonable treatment here. Indeed, the root and the progressive prefix both change as well according the subject’s person-numbers features even when non-contiguous with the agreement suffix (the ‘TNG ablaut’). The object prefix is given in Table 1; this is identical to the morphology used by nouns to reference possessors.
While the language is by and large head marking there is some case morphology, most significantly the ergative case marker -ma’. As is typical for the Papuan languages (Foley 2000) Yagarian ergativity is shallow and does not extend beyond case morphology. Moreover, beyond being obligatory for transitive subjects, for example (20a), it is also optionally available for intransitive subjects as in (20b) (Renck 1975:35).

(20) MOVE

a. hali-ma’ no-da-l-Ø-e
fire-ERG PROG-1SGU-burn-3SGA-IND
the fire is burning me (lo- entry in Renck 1977)

b. nala’a(-ma’) no’-Ø-e
who-ERG PROG-come-3SGA-IND
who’s coming? (Renck 1975:69)

From this we see that Yagaria is a fairly typical Papuan language of the Highlands. Also a typical part of Papuan languages are impersonals that also exist in Yagaria that will be described in next section.

6. Impersonals in Yagaria

Impersonals in Yagaria meet the general characteristics given in section 2. As can be seen from (21a) compared with (21b) the experiencer uses the encoding of objects: the same marker d(a)- agrees with the object in (21a) and the experiencer in (21b). The zero morpheme appears in both cases as the third singular agentive marker.

(21) MOVE

a. bali-ma’ no-da-l-Ø-e
fire-ERG PROG-1SGU-burn-3SGA-IND
the fire is burning me (lo- entry in Renck 1977)
All impersonals in Yagaria have these properties. Nonetheless, there are a number of subcategories to the supercategory of impersonal.

The vast majority of impersonals in Yagaria involve a nominal as in the above example (21b). This nominal is not the subject since it does not host the ergative suffix that is obligatory on transitive subjects (see section 5).\(^3\) The contrast between the nominals in (21a) and (21b) makes this obvious: the subject in (21a) (obligatorily) bears the marker -ma’ which is absent in (21b). Instead the nominal resembles a verb adjunct, part of a construction common in Yagaria (Windschuttel 2012:52-62).

Not all impersonals involve such a nominal. The example in (22) is an impersonal containing only a verb.

(22) MOVE

\[
\begin{array}{ll}
\text{dagaea} & \text{da-ha<no>k-Ø-e} \\
1\text{SG} & 1\text{SGU-feel.well<PROG>-3SGA-IND} \\
\end{array}
\]

I am feeling well (hago- entry in Renck 1977)

In a plurality of impersonals the experiencer agreement appears on the nominal instead of the verb as in the above examples. The example below in (23a) illustrates this—the same morpheme d- appears not on the verb but on the nominal, the verb adjunct. This is not unique to this construction and is the case with object agreement in many verb adjunct constructions (Windschuttel 2012, 2014). The transitive verb adjunct construction in (23b) employs this same marking on its verb adjunct in this case agreeing with the object.\(^4\) This variation

\[
\begin{array}{ll}
b. \text{agaso} & \text{no-d-ei-Ø-e} \\
\text{itchiness} & \text{PROG-1SGU-hit-3SGA-IND} \\
\end{array}
\]

I am itchy (agaso ei- entry in Renck 1977)

---

\(^{3}\)Indeed, in the optionally ergative languages of New Guinea inanimate subjects are those most likely to require marking where it is not generally obligatory on A arguments (see Donohue 2005).

\(^{4}\)Space will not allow properties of the verb adjuncts to be further examined (see Windschuttel 2012, 2014 for more details). One property that is worth mentioning is that they constitute a lexical word distinct from the verb itself. Examples such as (1) support this position: other words, here maloga, can intervene between a verb adjunct and the verb.

(1) YAGARIA

\[
\begin{array}{lll}
lap-oulaga & \text{maloga} & \text{ag-e-o} \\
1\text{PL-eye} & \text{here} & \text{look-PL-IMP} \\
\end{array}
\]

look at us! (oulago ago- entry in Renck 1977)
only reflects the same variation known in transitive predicates in Yagaria. For this reason it will not be of further significance to the discussion of impersonals in Yagaria.

(23) MOVE  
\[
\begin{align*}
  &a. \ d\text{-amota ei-d-i-e} \\
  &\text{1SG-fear hit-PST-3SGA-IND} \\
  &\text{I was afraid (-amota ei- entry in Renck 1977)} \\
  &b. \ d\text{-ageta eli-d-i-e} \\
  &\text{1SG-ear take-PST-3SGA-IND} \\
  &\text{he admonished me (-ageta eli- entry in Renck 1977)}
\end{align*}
\]

Impersonals in Yagaria fit into the general template given for impersonals in section 2. Neither of the optional parts of the definition are met by all forms of impersonals in Yagaria. The position of the experiencer is mostly irrelevant since it so rarely appears as an NP. A stimulus nominal is present for almost all Yagaria impersonals but is not necessary as demonstrated by the existence of a small number of simplex verbs like (22).

### 7. Non-canonical switch-reference in Yagaria

Switch-reference marking in Yagaria, as mentioned in section 3, is morphologically complex. In both DS and SS clause chains it employs markers that encode the features of the subject of the following clause. In SS clause chains this marker encodes the number-person features of the subject of both clauses. In DS clause chains it ‘anticipates’ the coming subject. The two markers are mostly identical in form and are presented in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Dual</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Person</td>
<td>-da</td>
<td>-ta’a</td>
<td>-ta</td>
</tr>
<tr>
<td>2. Person</td>
<td>-ka</td>
<td>-ta’a/-tati</td>
<td>-ta/-tapi</td>
</tr>
<tr>
<td>3. Person</td>
<td>-na/-ni</td>
<td>-da’a/-ti</td>
<td>-da/-pi</td>
</tr>
</tbody>
</table>

*Table 2. Paradigm of switch-reference agreement in Yagaria; the order is SS/AgtS/AntS where the two differ (Renck 1975:20)*

Schemata for the possible suffixes on the verb are given in (24) alongside examples of canonical switch-reference in (25). The suffixal morphology of a typical verb is summarised in (24a). As mentioned in section 4 this is site of
subject agreement (which follows a nominative-accusative system). This is contrasted with that of a DS medial verb in (24b) and SS medial verb in (24c).

(24)  
\[ a. \sqrt{V}\text{-TNS}\text{-AgrS}\text{-MOOD} \\
\[ b. \sqrt{V}\text{-TNS}\text{-AgrS}\text{-DS}\text{-AntS} \\
\[ c. \sqrt{V}\text{-SS}\text{-AgrS} \\
In DS medial verbs the mood slot of (24a) is replaced by an invariable DS morpheme and an anticipatory subject marker. The example (25a) gives a particular instance of this: the DS morpheme -(a)ga- fills the mood slot and is followed by a morpheme -da anticipating the first singular subject of the following clause. Thus a DS medial still bears normal agreement and tense inflection.

In (25b) we find a SS clause chain. The same third singular subject is shared throughout and this is marked by the morpheme -na on the medial verb. The verb itself is also substantially reduced; the SS morpheme is suffixed directly onto the root taking the place of tense and agreement morphemes as well as mood. This reflects the general form given in (24c).

(25) MOVE  
\[ a. \text{ba } d\text{-ami}\text{-d-an-aga}\text{-da} \\
\text{sweet.potato } 1SGU\text{-give-PST-2SGA-DS-1SG} \\
\text{no-d-o-e} \\
\text{PROG-eat-1SGA-IND} \\
\text{you gave me a sweet potato I am eating it (Renck 1975:115)} \\
\[ b. \text{dote'na } no\text{-do}\text{-na} \text{ ge } \text{hu\text{-d-i-e}} \\
\text{food PROG-eat-3SG.SS word do-PST-3SGA-IND} \\
\text{while he was eating he spoke (Renck 1975:108)} \\

However, as is the case in most Papuan languages this canonical use of switch-reference is not valid for all clause chains. In impersonals extra status is given to the experiencer: the DS-SS pattern of section 4.2 is followed. More significantly ‘clause skipping’ does exist.

In Yagaria the impersonals show a DS-SS pattern in their switch-reference as evidenced by (26) - (29). In (27), as is symbolically represented in (26), a personal clause precedes an impersonal clause and is marked for a different subject. This is to be expected since the verbs’ actor suffixes encode different features. In contrast to this when the impersonal and personal clause occur in the opposite
order, impersonal preceding personal, they marked as having the same subject as in (29). All similar examples from the texts in Haiman (1983) can be found in Appendix 1.

(26) [sbi: V-DS] [exp: …]  

(27) MOVE [boy: ol-u-ga-ni]  
garden take-1SGA-DS-3SG  
[da-hag:Ø-e]  
1SGU-be.well-3SGA-IND  
I am working and like it (hago- entry in Renck 1977)  

(28) [exp: V-SS] [sbi: …]  

(29) MOVE [gei-da da-hago-da]  
sickness-1SG 1SGU-recover-1SG.SS  
[bava’a bei-d-Ø-e]  
just.so live-PST-1SGA-IND  
I recovered and lived just so (Renck 1975:147)  

Although unmentioned, and likely undetected, by Haiman (1983) clause skipping exists in Yagaria (in the Hua dialect) and is present in his texts. There are two examples.

The first example involves a body part and is presented below as (31). The man is the subject of the first and third/fourth clauses while his penis is the subject of the second clause. Nonetheless, the first medial verb has same-subject marking, treating the man and his penis as the same. The third verb has different subject marking, as is expected, since there is change from the penis to the man. This fits into the clause skipping template repeated as (30) below.  

---

5 Lesley Stirling suggested, when this paper was presented, that this odd switch-reference could be because body parts and possessors are often treated as one for switch-reference (see Roberts 1997 for examples from Amele). This has not yet been observed for Yagaria and does not explain why the body part is only ‘raised’ for switch-reference between the first two clauses and treated as non-subject between the second and third clauses. If this is an example of clause-skipping, on the other hand, this is to be expected.
(30) [sbj] V-SS\(^1\) [sbj] V-DS\(^2\) [sbj]\(\ldots\)\(^3\)

(31) HUA \(\ldots bavi-na\)\(^1\)
go.up-3SG.ss
\[gna-giva-mo\] \(\ldots na\)
penis-great-DIS thus
\[gveda\] \(\ldots na-mo\)
rope like thing-DIS be-3SG-COND-DS-3SG
\[rgavaurgavaibnu-na\] \(\ldots ku\-'a\-'vi\) \(\ldots airo-na\)\(\ldots\)\(^3\)
coil.up-3SG.ss bag-3SG-IN put-PFV-3SG.ss
he went up and his penis, a great rope like object, he had coiled up keeping it in his bag (Haiman 1980:484)

The other example (32) skips over a woman in a chain of clauses describing the actions of the male protagonist. This man is the subject of the first clause and resumes this role at the end of the excerpt. The woman speaks in the middle section and her question is repeated twice, each utterance followed by \(huna\) which is acting here as a quotative (it has its source as a SS medial verb; thus it is unclear if there is one clause or three; this, however, does not affect the clause skipping analysis). The verb preceding the speech is marked same-subject despite the man arriving and the woman speaking being distinct. The distinction between these two subjects is acknowledged by the last verb in the example through its DS morphology. The pattern SS-DS is the same clause skipping pattern as above.
(32) HUA [vzahu-na-bi’-mo] 
arrive-3SG.SS-BEN-DIS 
“zahu-r-a-ne” bu-na 
do.what-PFV-2SGA-INT say-3SG.SS 
a’-ma-mu’ “zahu-r-a-ne” bu-na[1] 
that-woman-ERG do.what-PFV-2SGA-INT say-3SG.SS 
[b-i-ga-na-bi’-mo][2] 
say-3SGA-DS-3SG-BEN-DIS 
[ma dea…][3] 
that man 
when he arrived, the woman asked him, “What are you doing?” and the man… lit. he arrived and saying ‘what are you doing’, the woman saying ‘what are you doing’ she said, and the man…
(Haiman 1980:490-1)

These examples make sense with our knowledge of the languages of the Highlands. The unusual switch-reference follows this same pattern described for other Papuan languages in section 3.1.

Thus Yagaria has both clause skipping and special switch-reference in impersonal clauses. However, the two are distinct from one another. This is clear from the opposite order of the DS and SS marking in (33) and (34).

(33) [sbj] V-SS][1] [sbj] V-DS][2] [sbj] …][3]
(34) [sbj] V-DS][1] [exp] V-SS][2] [sbj] …][3]

There is one example of switch-reference around an impersonal clause in the opposite order, (36) i.e. the same pattern in clause-skipping. However, this example is unusual because there is a switch-reference marker on the second verb though no clause follows. In clause skipping switch-reference marking must both precede and follow the skipped clause as in (35). Thus this is likely an example of clause skipping with the obligatory two instances of switch-reference (but no clause following the second). It just so happens that the verb involved in this case is impersonal.

(35) [sbj] V-SS][1] [sbj] V-DS][2] [(sbj) …][3]
(36) HUA “… [bu-ro-na][1] [aipa ro-re-Ø-ga-na][2] [Ø][3] say-PFV-3SG.SS guts burn-PFV-3SGA-DS-3SG …” he said and was angry (Haiman 1980:503)
These data would suggest Roberts (1997) overgeneralised when suggesting that special discourse uses of the switch-reference system could be a full explanation of switch-reference in impersonal clauses in all Papuan languages. Clause skipping and impersonal switch-reference are not always part of the same pattern. Yagaria uses the expected SS-DS pattern for clause skipping, but for impersonal switch-reference a distinct DS-SS pattern is employed.

Thus in Yagaria impersonal clauses are not simply skipped but some other process is at work causing their particular switch-reference.

8. The subject in Yagaria

This discussion of switch-reference will inform the understanding of the subject in these clauses. There are three possibilities for the subject. The third singular person morphology could be the subject, there could be no subject or the experiencer could be the subject. Case morphology is of no assistance since none of the nominals carry ergative morphology.

Agreement morphology points to a third singular subject. The suffixal agreement on impersonal verbs is regularly used for incontrovertible subjects. However, the data from switch-reference dispute this. In the following example one can see the exceptional switch-reference in an impersonal-personal clause chain. As a SS medial verb the impersonal verb has no subject agreement since this slot is filled by the SS morpheme. Normally, this lack of subject agreement does not matter since the SS morpheme marks this same argument; however, in this case the SS morpheme references the experiencer and not reference a third singular actor. Since the agreement morpheme was the only reference to a third singular participant and it does not appear here one may conclude that it does not represent an argument at all but was rather a default, filling a slot. Thus the third singular agreement morpheme is not the subject.

(37) MOVE

\[
\begin{align*}
&\text{sickness-1SG} & \text{1SGU-recover-1SG.SS} \\
&\text{just.so} & \text{live-PST-1SGA-IND} \\
\end{align*}
\]

I recovered and lived just so (Renck 1975:147)
Two possibilities remain: no subject or a non-canonically marked experiencer subject. Again the only evidence either way is the impersonal predicate’s behaviour in clause chains. Switch-reference in examples, such as (37), references the experiencer (in the opposite direction it does not; the reasons for this are not clear). Switch-reference is known to follow roles apart from the subject in clause skipping but as section 7 explained this is not clause skipping. Hence switch-reference is following its canonical function of tracking subjects.

This suggests that the experiencer is the subject albeit a non-canonically marked one. This brings these constructions into line with other non-canonical subjects, most famously Icelandic albeit defined by agreement morphology rather than case (Andrews 1982, 2001 among many others).

9. Object experiencer

The behaviour of impersonals in clause chains betrays a deeper complexity. As explained in section 8 the experiencer is sometimes treated as the subject. In other cases the same test can treat this same argument as if it were not the subject. This section will discuss this and propose an analysis for the construction in these cases.

The example (38) shows this contradictory behaviour of impersonals in clause-chaining. This is identical to example (37) above except that the different-subject marking is employed between the two clauses instead of same-subject marking. This alternative grammatical packaging is a general possibility albeit a substantially less frequent one (Renck 1975:147).

(38) MOVE  
\[ \text{sickness-1SG } da-bago-e-ga-da^1 \]  
\[ \text{1SGU-recover-3SGA-DS-1SG} \]  
\[ \text{just.so } bei-d-u-e^2 \]  
\[ \text{live-PST-1SGA-IND} \]  
I recovered and lived just so (Renck 1975:147)

This was the only significant evidence for the subjecthood of the experiencer. Hence there is no evidence of subjecthood in this case. Instead it is likely that, in these marginal cases, the experiencer corresponds to the canonical function marked by undergoer agreement, object.
There are a number of possible analyses of the syntax where the experiencer is object viz. the subject may be the third singular or there may be none. However, for lack of decisive data either way this will not be examined further.

Thus two competing analyses exist in the grammar of speakers: the experiencer in these constructions is an oblique subject or it is an object befitting its morphological expression. The next section will propose an answer to the question whether and how these two grammars are related to one another diachronically.

10. History of the construction
There is some evidence to suggest that the two possible syntactic interpretations of the impersonal construction are in a diachronic relationship with one another. Namely, the subject experiencer interpretation developed out of the older object experiencer construction.

This explains why the agreement morphology used by the experiencer is shared with that of regular objects: the experiencer was an object earlier in the language and hence this marking was required. Furthermore, the typical assumption is that these constructions normally involve actual objects: this is implied by the Evans’ (2004) renaming them ‘object’ experiencer constructions and made explicit in Cole and others’ (1980) study of similar constructions around the world.

Cole and others (1980) study the acquisition of subjecthood in languages from three unrelated families. They conclude that objects become subjects in a particular way: behavioural, purely syntactic properties accrue first before later morphological ‘coding’ properties are added making the subject indistinguishable from any other.

The gain of coding properties by ‘oblique’ subjects, that is, apparent objects with subject behavioural properties, is well attested. We have textual witness to this change as it occurred in the history of English (Allen 1995). This is the change that their non-Indo-European examples supported as well.

The earlier change from object to oblique subject was only evidenced by Germanic in their paper. Their analysis of the relevant Gothic data has been
questioned and the possibility has been raised that oblique subjects are stable and were present in Proto-Germanic (e.g. Barðdal & Eythórsson 2009, Eythórsson & Barðdal 2005).

This change, from true object to oblique subject, is that that this section will examine for Yagaria. This draws on outside data from closely related languages, but the most telling evidence will be language internal. This will rely on the preservation of historical forms in an idiom. By definition idioms are marginal and hence the conclusions rely on only marginal data. This idiom involves a serial verb construction and thus an understanding of the origin of these constructions in the language will be necessary.

### 10.1 History of the SVC in Yagaria

Serial verb constructions (SVCs) in Yagaria derive historically from clause chains. This is most clearly evident in their use of the same morphology. This is in line with the generality of Givón (2001:51) for the historical source of SVCs in all languages. The use of this switch-reference morphology in SVCs often differs from that known in clause chains and this is the clearest identifier that a clause chain has been reanalysed as an SVC.

The switch-reference markers often appear in apparently anomalous sequences in examples such as (39). Were this series of verbs to be read as a clause chain its meaning would be semantically unlikely and, more significantly, it would differ from that given in translations. The clause chain would mean ‘I shoot you with an arrow and I fall over and put you’. The construction instead has a more sensible meaning ‘I shoot you and cause you to fall down’. This translation makes sense if we recognise that this construction has been reanalysed as an SVC.

(39) **MOVE**

\[
\begin{align*}
\text{bali-loti’} & \quad \text{ga-bao-da} \\
\text{arrow-INS} & \quad 2\text{SGU-shoot-1SG.SS} \\
\text{gagupao-da} & \quad \text{ga-to-g-u-e} \\
\text{fall.over-1SG.SS} & \quad 2\text{SGU-put-FUT-1SGA-IND}
\end{align*}
\]

I shall shoot you down with an arrow (Renck 1975:144)

If we take other examples into account we can observe that in SVCs only same subject markers ever appear and these encode the subject of the whole SVC
(SVCs in Yagaria only ever have one subject throughout; see Windschuttel 2012:42-45). This is true of (39) where all SS markings refer to the first singular subject, even on the verb, gapupao- ‘fall over’. This is despite the fact the ‘faller’ is actually the second singular object.

A similar line of reasoning will be followed in the next section in examining a particular idiom containing switch-reference markers. These markers do not admit a possible clause chain interpretation but may result from the historical creation of an SVC.

10.2 Idiom containing a frozen impersonal SVC

As explained in the previous section SVCs in Yagaria derive historically from clause chains. This knowledge we can apply to analysing a particularly unusual SVC containing an impersonal, (40).

(40) MOVE ao-da halupao-na da-ba-Ø-e
tread-1SG.SS slippery-3SG.SS 1SGU-shoot-3SGA-IND
I am slipping (halupao- hao- entry in Renck 1977)

This string of verbs contains a possible impersonal, the experiencer is referenced by a prefix on the last verb hao- which apparently references a third singular subject. Of more interest are the preceding verbs and their agreement features. The morpheme -da on the first verb is first singular while -na on the second verb is third singular. Thus in this SVC there is a change in the person of the SS morphemes. A string of same subject morphemes with necessarily differing referents is otherwise unknown in the language.

This may make sense from a historical standpoint, if the subject has been different at different points of time. There have been two stages of development in the SVC in question and the subject has changed between the stages (as has already been supposed to have happened in all impersonals).

As stated in the previous section SVCs always use SS morphemes that reference the SVC’s subject. At some point in the past an SVC was formed out of the two verbs halupao- and hao-. At that time the subject of the impersonal clause was third
singular. This inner SVC is still present in the contemporary language, for example in (41) below.

(41) MOVE  
\begin{tabular}{lll}
\vspace{-0.7cm}
\textit{igopa} & \textit{halupao-na} & \textit{da-ha-O-e} \\
ground & slippery-3SG.SS & 1SGU-shoot-3SGA-IND \\
\end{tabular} 
\begin{tabular}{lll}
\vspace{-0.7cm}
& I slipped on the ground (halupao- hao- entry in Renck 1977) \\
\end{tabular}

At a later time the experiencer was reanalysed as the subject. The third singular SS morpheme on \textit{halupao-} had now become non-referential; this was likely the reason it was preserved. After this reanalysis the SVC in (41) was formed; the first singular SS morpheme on the first verb agrees with the new experiencer subject. Hence, the contradictory sequence of SS morphemes that we now see.

Since this historical analysis is predicated on the historical antecedence of the object experiencer construction, if it is correct, this gives a small piece of evidence that the experiencer was originally the object and only later reanalysed as the subject.

This fits with the situation we find in the languages most closely related to Yagaria where the object experiencer is the norm. Kamano will serve as suitable representative of the Kamano-Yagaria language family and will date this change after the breakup of the family.

\textbf{10.3 Impersonals outside of Yagaria}

The greater time depth of the object experiencer pattern is corroborated by data from Kamano where we only find object experiencers in impersonals.

Yagaria belongs to the Kainantu-Gorokan family (a possible primary branch of the TNG family). Yagaria is on the Goroka side of this family and more specifically Kamano-Yagaria (Foley 1986:258). Even among these ~6 languages, there are languages in which switch-reference treats the impersonal experiencer consistently as object in contrast to what has already been established for Yagaria.

This is the situation in Kamano. Experiencers in impersonal clauses and subjects in personal clauses may have the same referent across a clause chain. Such a chain is always marked for change of subject (in the sentences examined in the Kamano
dictionary, Drew & Payne 2005). It does not matter in which order the clauses are.

The following examples in (42) display the two orders possible. One will note in both cases the morpheme -ge- is used, cognate with Yagaria -ga-, indicating a change in subject. The subject and experiencer are both first person singular: this is clear from the first singular marking on the verb adjunct and final verb in (42a) while in (42b) the medial verb and object prefix show agreement for these features. More Kamano examples of switch-reference involving impersonal clauses are given in Appendix 2.

(42) KAMANO  

a. Nga' ategeno' nvaza neroe  
\[ n\text{-}ga'\quad ata\text{-}i\text{-}ge\text{-}no' \] \[ n\text{-}vaza\quad ne\text{-}re\text{-}u\text{-}e \]  
1SG-hunger 1SG-yawn  
\[ \text{hunger-3SGA-DS-3SG} \quad \text{PROG-stab-1SGA-EMPH} \]  
I am hungry and I yawn (Drew & Payne 2005:22)

b. Eri'za eri vava hu'nogeno' ksro' nhe'nee  
\[ eri'za\quad eri\text{-}vava\quad bu'ne\text{-}u\text{-}ge\text{-}no' \] \[ ksro'\quad n\text{-}be\text{'ne\text{-}i\text{-}e} \]  
work tired 1SGU-do-PST-3SGA-EMPH  
\[ \text{work-always do-PST-1SGA-DS-3SG} \quad \text{1SGU-do-PST-3SGA-EMPH} \]  
I worked so hard all the time I was tired (Drew & Payne 2005:55)

The subject interpretation is not the norm even among the most closely related languages. Kamano evidences this above. Section 10.2 suggested that the interpretation with the experiencer as object was older than as subject in Yagaria. These data from Kamano accord with this; non-canonical experiencer subjects cannot be traced back even to the smallest family containing Yagaria. This implies that this change from object to oblique subject was a relatively recent one.

11. Conclusion

The conclusions of this paper are twofold. In one Papuan language, Yagaria, impersonal constructions involve non-canonically marked subjects on par with those of Icelandic and others. The other is that this construction derives from one in which the experiencer was the object. Both of these conclusions are quite
tentative relying on a small set of data points but appear to be the best fit for the current data.

On the way to the first conclusion an important result was made about the switch-reference system of Yagaria and the Papuan languages in general. Clause skipping of the form known from Longacre (1972) is present in Yagaria. This is in direct contradiction to the prediction given by Roberts (1997) that impersonal switch-reference is purely pragmatic and that pragmatic switch-reference exists in the two forms in (43) and (44) below.

\[(43) \quad [\text{sbji} \ V-\text{SS}]^1 \ [\text{sbji} \ V-\text{DS}]^2 [\text{sbji} \ \ldots]^3\]
\[(44) \quad *[\text{sbji} \ V-\text{DS}]^1 \ [\text{sbji} \ V-\text{SS}]^2 \ [\text{sbji} \ \ldots]^3\]

Instead I contend pragmatic switch-reference exists in only the first form. The switch-reference in impersonals in those languages which follow a DS-SS pattern is something different. Indeed, at least in the Yagarian case, it is a direct reflection of the subjecthood of the experiencer. This brings us to one of the main conclusions of this paper. Rarely does the subject vary from the semantic role of actor in Papuan languages. In Yagar ia, however, it does: specifically in impersonals the experiencer is the subject although it did not cause nor does it have any control over the experience. That the subject should have any distinct existence from the actor/cause is significant to the study of the argument structures of Papuan languages.

There is some indication that this oblique subject in Yagaria is a recent innovation and that the experiencer as non-controller was previously subsumed to object. Section 10 presents argumentation for the historical precedence of the experiencer object analysis. This provides data for the theory of Cole and others (1980) governing the acquisition of subject properties by experiencer objects. The proposed existence of a change from object to non-canonically marked subject has been controversial: some evidence has been gathered against its attestation in Germanic (e.g. Barðdal & Eythórsson 2009, Eythórsson & Barðdal 2005). It is this pathway that is possibly validated by Yagaria, a language totally unrelated, both typologically and genetically, to those which Cole and others (1980) studied.
Abbreviations
A actor agreement
CONT continuous
DIS discourse marker
EMPH emphatic mood
FUT future
INCONS inconsequential
LIM limiter enclitic
NDU non-dual
PURP purposive
TODP today’s past
U undergoer agreement

Acknowledgements
This paper is based on research for an Honours thesis written under the supervision of William Foley and Jason Johnston at the University of Sydney.

References
Andrews Avery 2001 Non-canonical A/S marking in Icelandic In Alexandra Aikhenvald, Robert Dixon & Masayuki Onishi (eds) Non-canonical marking of subjects and objects, 85-111 Amsterdam: John Benjamins
Austin Peter 1981 Switch-Reference in Australia Language 57:309-334
Barðdal Jóhanna & Thórhallur Eythórsson 2009 The Origin of the Oblique Subject Construction: An Indo-European Comparison In Vit Bubenik, John Hewson & Sarah Rose (eds) Grammatical Change in Indo-European Languages 179–193 Amsterdam: John Benjamins

Dowty David 1991 Thematic proto-roles and argument selection Language 67:547-619

Evans Nicholas 2004 Experiencer objects in Iwaidjan languages (Australia) In Peri Bhaskararao & Karumuri Subbarao (eds) Non-nominate subjects Vol. 1 169–192 Amsterdam: John Benjamins

Evans Nicholas 2011 Valence in Nen Presented at Valence Workshop Leipzig MPI-EVA April 2011

Foley William 1986 The Papuan Languages of New Guinea Cambridge: Cambridge University Press


Givón T 1997 Grammatical relations: An introduction In T Givón (ed) Grammatical Relations: A Functionalist Perspective 1–85 Amsterdam: John Benjamins

Haiman John 1980 Hua: A Papuan language of the Eastern Highlands of New Guinea Amsterdam: John Benjamins

Haiman John 1991 Hua-English dictionary Wiesbaden: Otto Harrassowitz

Haiman John & Pamela Munro (eds) 1983 Switch Reference and Universal Grammar Amsterdam: John Benjamins


Pawley Andrew 2006 On the argument structure of phrasal predicates in Kalam, a language of the Trans New Guinea family Paper presented at The 32nd Annual Meeting of the Berkeley Linguistics Society, Berkeley, California, USA


Appendix 1: Impersonal examples from Hua texts (Haiman 1980)

(1)  ...\text{-}ga-na\text{ }na-vi'\text{ }bgota'\text{ }bai\text{-}O\text{-}ma'\text{ }de-mo
...-DS-3SG\text{ }thing\text{-}IN\text{ }one\text{ }be\text{-}3SGA\text{-}GEN\text{ }man\text{-}DIS
na\text{-}bu-na\text{ }siki\text{-}na-mo\text{ }ba<\text{O}>\text{o-na}\text{ }nana-mo
thus\text{-}do\text{-}3SG.SS\text{ }sick\text{-}thing\text{-}DIS\text{ }afflict<3SGU\text{-}3SG.SS\text{ }sickness\text{-}DIS
ba<\text{O}>\text{o-na}\text{ }bu-re-ga-na\text{ }fri\text{-}O\text{-}si\text{-}mi'
afflict<3SGU\text{-}3SG.SS\text{ }do\text{-}PFV\text{-}DS\text{-}3SG\text{ }die\text{-}3SGA\text{-}SBJV\text{-}PURP
bu\text{-}hil\text{-}O\text{-}gi\text{-}te\text{ }^6\text{ }do\text{-}PROG\text{-}3SGA\text{-}DS\text{-}3PL.
...and this unmarried man was afflicted with sickness to the point of death. (Haiman 1980:481)

^6 V\text{-}SBJV\text{-}PURP\text{ }bu\text{-}\text{ is a construction meaning ‘to be about to V’ (Haiman 1980:179)}
(2)  ... mabo'-mo  d-gorai-ka  ri-ka  u-ka
d-mi-bai-pa'  na-mo  do-bau-ga-na-bi'-mo
d-mupai-re-Ø-ga-da  d-mupa-vina  ge  bar-u-e
this-DIS  1SGU-fool-2SG.SS  leave-2SG.SS  go-2SG.SS
1SGU-give-PROG.2SGA-GEN  thing-DIS  eat-PROG.1SGA-DS-3SG-BEN-DIS
1SGU-full-PFV-3SGA-DS-1SG  1SG-belly-IN  word  hear-1SGA-IND
So that's how you've been tricking me; I've been eating what you're giving me and feeling full and having belly ache. (Haiman 1980:487)

(3)  ... na  auva-roga  ge-mana  ha<Ø>ago-bai-Ø-ga-na
na-Ø-frame  poor easy do-3SGA-DS-3SG
...that body-ALL see-INCONS feel.good<3SGU>-CONT-3SGA-DS-3SG
rgi'  'a'-rmi-si-Ø-ma'  aipa  kri-Ø-e
really NEG-go.down-FUT-3SGA-GEN guts  grow-3SGA-IND
...looking at her body, he really felt good and he really didn't feel like going down (Haiman 1980:492)


(1)  zn-avesra  h-i-ge-'ne  kho  hu-ku'  a'-e
3DU-tired do-3SGA-DS-3DU rest do-PURP come.3DUA-EMPH
those two are tired and have come to rest

(2)  Zferina'a  omne  amne  h-i-ge-no  mse'-ne-Ø-e
frame  poor easy do-3SGA-DS-3SG lie-PAST-38GA-EMPH
he is tired and is lying down

(3)  Ø-avesra  hi-ge-no'  mni  fru  hu'-ne-Ø-e
3SG-tired do-3SGA-DS-3SG rest calm do-PST-3SGA-EMPH
he was tired so he rested
The Virtual Atoll Task:
A spatial language elicitation tool

Jonathon Lum and Jonathan Schlossberg
Monash University and The University of Newcastle
jlum2@student.monash.edu
jonathan.schlossberg@uon.edu.au

Abstract. This paper describes and evaluates the Virtual Atoll Task (VAT), a new elicitation tool for research on spatial language. The task, which elicits spatial descriptions as participants direct each other through a simulated landscape, provides a more realistic, first-person perspective in comparison with previous tabletop space games (e.g., Levinson et al. 1992; Senft 1994; Wilkins 1993), and aims to prompt large scale spatial descriptions that may be uncommon in tabletop space. We discuss the methodology of the task then comment on its strengths and limitations as an elicitation tool, reporting on pilot studies in Jaluit Atoll (Marshall Islands) and Laamu Atoll (Maldives). We conclude by discussing the potential for the VAT to be modified for use in other topographic environments, as well as for the task to be employed for the investigation of other domains.

Keywords: virtual environments, spatial language, elicitation, linguistic fieldwork, frames of reference
1. Introduction

Conventional linguistic fieldwork has tended to prioritise the collection of texts produced naturally within the community, such as conversations or traditional stories. However, it has been recognised that research into specific linguistic domains can often benefit from a more targeted approach to elicitation, through the use of elicitation tasks or games with specially designed stimuli. One domain for which this has particularly been the case is the domain of spatial language. Researchers from the Language and Cognition Group at the Max Planck Institute for Psycholinguistics, Nijmegen, have developed a suite of elicitation tasks for exploring various aspects of spatial language. These include ‘route description’ games (Wilkins, 1993), photo-photo matching games (Levinson et al., 1992), photo-object matching games (Senft, 1994), and others. Despite certain shortcomings (see below), such tools are a convenient way of eliciting language rich in spatial terms, and when supplemented by more naturalistic texts, can help to provide valuable insights into the dominant systems of spatial reference in a language.

In this paper we report on the ‘Virtual Atoll Task’ (VAT), a newly developed tool for the elicitation of spatial language. The VAT is a computer game set in a virtual atoll environment (see Fig. 1 below). The game and the virtual environment in which it takes place were designed by the authors using Unity, a program for computer game development. The VAT is played by two native speaker consultants who cooperate to search for a number of hidden treasure chests around the atoll. It aims to elicit highly precise spatial descriptions as one player – to whom the locations of the treasure chests have already been revealed – verbally guides his partner through the game.

The VAT provides a controlled way to observe how speakers give directions and locative descriptions in a large scale environment (albeit a virtual one), and thereby gets around many of the difficulties associated with the use of pre-existing tabletop games, as well as the challenges associated with directly observing speakers as they move through actual large scale space. Although it is not

---

1 See http://fieldmanuscripts.mpi.nl.
intended to replace more established stimuli, we submit that the VAT is a useful addition to the suite of elicitation tools already developed for investigating spatial language.\textsuperscript{2}

\textit{Figure 1. Screenshot from the Virtual Atoll Task.}

This paper is divided into five sections. Section 2 provides some background to the cross-linguistic study of spatial language and to some of the existing elicitation methodologies in this domain. Palmer’s (2015) ‘Topographic Correspondence Hypothesis’ is also introduced in order to explain the particular interest in atolls. Section 3 describes the methodology of the VAT as well as the design of the virtual environment in the game. Section 4 discusses some of the linguistic findings from the use of the VAT at our field sites in Laamu Atoll (Maldives) and Jaluit Atoll (Marshall Islands), and evaluates the strengths and weaknesses of the VAT as an elicitation tool, in comparison to other spatial language elicitation tasks. Finally, Section 5 concludes by summarising the features of the VAT and by discussing other potential applications for the task. In particular, we note that while few researchers are likely to employ the task in atoll-based communities, a

\textsuperscript{2} Readers wishing to use the VAT in their own research are encouraged to contact the authors.
larger number may be interested in modifying the virtual environment to alternative topographies, or may wish to use the game for different purposes entirely.

2. Spatial language and its elicitation

2.1 Background

Talking about space is a fundamental part of speaking a language. Every day, we find ourselves in situations where we must describe the locations of objects or people, or talk about which way they are facing or where they are going. Naturally, linguists have wished to study how languages encode spatial descriptions, and to discover the level of cross-linguistic variation or uniformity that exists in terms of both grammatical structures and semantic categories. Several subdomains of spatial language have been found to vary across languages, including deixis, topological relations, motion descriptions, and frames of reference (Levinson & Wilkins, 2006). In addition, linguistic variation in the spatial domain correlates with and may even be a cause of variation in non-linguistic spatial cognition, as reflected in spatial memory and reasoning (Levinson, 2003; e.g., Levinson, Kita, Haun, & Rasch, 2002; Majid, Bowerman, Kita, Haun, & Levinson, 2004; etc.). The study of spatial language is therefore important not only to linguistics, but to our understanding of how different cultures conceptualise physical space.

2.2 Eliciting spatial language

Most grammars of languages include some information on structures that express spatial concepts, and many linguists have published more detailed accounts of the spatial systems in particular languages, based on extensive fieldwork (e.g., Levinson & Wilkins, 2006; Senft, 1997). How is such information obtained when in the field? To some extent, researchers have relied on direct observation, immersion in the language and culture, and the collection and close analysis of texts such as narratives and conversations. However, some have also used more targeted elicitation tasks involving carefully developed stimuli. Although
somewhat less naturalistic, the use of such tools has a number of important advantages: (i) the controlled nature of the tasks allows for cross-linguistic comparisons to be made more soundly; (ii) researchers can easily manipulate variables of interest, such as the size, shape or layout of the stimuli, or the direction the participants are facing; and (iii) tasks can be designed to elicit texts rich in highly precise spatial terms, whereas more naturalistic texts might feature such terms only occasionally.

For certain areas of spatial language, many researchers have found it beneficial to elicit instructional texts produced by pairs of native speaker consultants. This was the approach favoured by the MPI’s Language and Cognition Group during their ‘space’ project in the 1990s and early 2000s, and the basic methodology has subsequently been taken up by others including Bennardo (2000), Li and Gleitman (2002) and Dasen and Mishra (2010). The methodology generally involves sitting two consultants side by side but separated by a screen. One consultant acts as ‘director’ and must describe the stimuli in front of her to the ‘matcher’, who must try to select or manipulate the corresponding stimuli on her own side. The matcher is usually allowed to ask questions.

Stimuli are carefully designed to maximise the need for consultants to employ highly precise spatial language during the task. For example, in the ‘Man and Tree’ game (Levinson et al., 1992), the director cannot simply say, ‘in this photo there is a man and a tree’, since the stimuli include many photos showing a man and a tree in various configurations. Therefore, the game is useful for eliciting more precise spatial descriptions such as ‘the man is in front of the tree’ or ‘the man is north of the tree’. Man and Tree is an example of a ‘photo-photo’ matching game; other games involve ‘photo-object’ matching (e.g., Senft, 1994) or ‘object-object’ matching, which includes ‘route description’ games (e.g., Wilkins, 1993).

5 The MPI’s methodology was inspired by earlier experiments in psychology of language, such as the Human Communication Research Centre Map Task (Anderson et al., 1991; http://groups.inf.ed.ac.uk/maptask/maptask-description.html).

4 Other configurations may be used as variants.

5 Later variants have used photos of a different (toy) man and tree (e.g., Terrill & Burenhult, 2008), or have used photos of different objects entirely, such as a girl and an umbrella (Li & Gleitman, 2002) or a ball and chair (Bohnemeyer, 2011; Polian & Bohnemeyer, 2011).
Despite the ability of these games to elicit precise spatial relations, there are certain limitations and shortcomings. For one thing, the particular design of the task affects the results. For example, if a set of photos includes a lot of symmetrical objects, one would expect to see fewer tokens of intrinsic frame of reference (e.g., ‘the tree is left of the man’, i.e., at the man’s own left side) than if the stimuli included mostly faceted objects. Another issue is that because of the nature of the task, consultants might resort to strategies that are linguistically available but not common in their language. Claims about the prevalence of certain spatial strategies in a language should therefore not be based on director-matcher style tasks alone. And of course, there are typically a range of practical challenges associated with running such tasks in the field, including consultants being unfamiliar with certain stimuli or not understanding what to do in the game. To an extent, however, these difficulties can be mitigated by collecting data from multiple elicitation tasks in combination with more naturalistic texts, so that one’s findings do not depend too much on the design of any particular task.

A different limitation (and one that is of special interest to this paper) is the small scale, tabletop nature of most director-matcher games. In many languages, the linguistic strategies preferred in small scale, manipulable space are not the same as those used when referring to the wider world. English, for example, tends to reserve cardinal directions like north and south for large scale space, such as when describing regions of countries or long distance voyages. Tabletop director-matcher games like Man and Tree are poorly suited to eliciting vocabulary that is largely restricted to those kinds of contexts.

On the other hand, it can be problematic to observe people using spatial language as they navigate through large scale space – aside from the logistical challenge of getting quality recordings when exposed to the elements and noisy vehicle engines, one may well find that speakers use precise spatial language infrequently. This may be because their communicative needs are adequately served by pointing gestures and the use of place names (e.g., for the destination), or because they are travelling along a route they have taken many times before and so spatial descriptions are unnecessary. The VAT was specifically designed to provide a method for eliciting large scale spatial descriptions while avoiding some of the
issues associated with existing director-matcher tasks and also those associated with direct observation of real world way-finding.

2.3 Spatial language and the environment: the case of atolls

Before describing the methodology of the VAT in Section 3, it is first necessary to introduce the theoretical background and hypotheses which inspired the task. This will also address why the VAT was designed specifically for atoll environments (though as we explain in Section 5, it has the potential to be modified for other topographies).

Palmer (2015) proposes a ‘Topographic Correspondence Hypothesis’, predicting that languages spoken in similar topographic environments should exhibit similar subtypes of the absolute frame of reference in accordance with local topography, while languages spoken in different topographic environments should be expected to have diverse systems of absolute frame of reference. Although other researchers have discussed the relationship between frames of reference and the environment, Palmer’s hypothesis relates specifically to the topography of the language locus, rather than to an indoors/outdoors variable (e.g., Levinson et al., 2002; Li & Gleitman, 2002) or to the ‘ecological zones’ (temperate, tropical, etc.) or ‘subsistence modes’ (hunter-gatherer, stable agriculture, etc.) considered by Majid et al. (2004). It is also more fine-grained in that it makes predictions as to particular subtypes of the absolute frame of reference (e.g., cardinal systems, landward-seaward systems, uphill-downhill systems, etc.), rather than claiming that absolute frame of reference as a whole correlates with some environmental feature or other.

Palmer (2015) observes that the spatial systems in a number of languages appear to support the Topographic Correspondence Hypothesis (see also: Palmer, 2002, 2004, 2007). For example, Aralle-Tabulahan (Austronesian), Samo (Papuan), Dyirbal (Australian) and Florutz German (Indo-European) are all spoken in environments featuring mountains and rivers, and all four languages possess similar absolute directional systems that include terms for ‘uphill’, ‘downhill’, ‘upriver’, ‘downriver’ and ‘across river/valley’. However, to test the hypothesis in
a wider range of languages and topographic environments, Palmer (2015) proposes an ‘Environment Variable Method’ with two dimensions of comparison: first, the language is held constant and the environment varied (i.e., observe speakers of a single language or of closely related languages that are spoken in different topographic environments); and second, the environment is held constant and the language is varied (i.e., observe speakers of unrelated languages that are spoken in similar topographic environments).

Palmer identifies atoll environments as a useful test case for the Topographic Correspondence Hypothesis and Environment Variable Method. An atoll is a ring of coral reef and flat, narrow strips of land surrounding a central lagoon (see Fig 2 below).

*Figure 2. Satellite photograph of Atafu Atoll, Tokelau.*

The lagoon on the inside of the atoll and the ocean on the outside play very different roles in the lives of atoll dwellers, and both are visible from many points on land. The locations of the two bodies of water are therefore highly salient features of this rather unusual environment. The Topographic Correspondence

---

Hypothesis predicts that the spatial systems of atoll-based languages should attend to this distinction between ocean and lagoon, and indeed the atoll-based languages of the Pacific do employ ‘lagoonward-oceanward’ axes (Palmer, 2007; 2015). However, as Palmer observes, these languages are all members of the Oceanic branch of Austronesian, and so are not ideal candidates for comparison. Dhivehi, an Indo-European language spoken throughout the atolls of the Maldives in the Indian Ocean, is a more suitable language for the second dimension of the Environment Variable Method. As for the first dimension, Marshallese migrant communities in the United States provide a useful comparison to their relatives still living on atolls in the Marshall Islands.

In order to test the Topographic Correspondence Hypothesis on atoll-based languages, we have collected spatial language data from both Dhivehi and Marshallese, with the latter subdivided into atoll-based Marshallese and US Marshallese. To this end, we have employed a range of spatial language elicitation tools in tandem with more conventional language documentation techniques. The VAT is but one of these tools, though it is one that has a special place because it requires participants to navigate through a version of their environment. It thereby provides a window on how different atoll communities speak about and conceptualize the topography around them, which has important implications for the Topographic Correspondence Hypothesis and for our more general understanding of the interaction between language, culture, cognition and the environment.

3. Methodology and game design

3.1 Procedure

The VAT is a cooperative computer game played by two consultants, a ‘director’ and a ‘matcher’. The role of the director is to guide the matcher to a number of treasure chests which are hidden in various locations around the virtual atoll. The

---

We use the terms ‘director’ and ‘matcher’ as these are in keeping with earlier space games (see §2.2), even though the methodology is somewhat different.

~ 90 ~
The director does not interact with the controls in any way. The role of the matcher is to operate the controls and follow the director’s instructions in order to discover and open the chests. Both players are encouraged to speak to each other throughout the task, but are asked not to point at the screen. The basic procedure is as follows, and a photograph of two Marshallese men playing the VAT is provided in Figure 3:

1. The director watches a short video showing the locations of the treasure chests while the matcher waits outside the room. The video is actually a screen capture recording of the first author playing the game without any director and accompanied by some background music.

2. The director is invited to watch the video again and to play the game himself, until familiar with the locations of the chests. During this process, the researcher may intervene to assist with any issues relating to the controls, but should not prime the director by using any spatial vocabulary or by offering strategies for remembering the locations of the chests.

3. The matcher is invited into the room and given a small amount of practice with the controls.

4. When both players are ready, the researcher loads the game scene and turns on any recording equipment (ideally, a good quality audio recorder connected to two lapel microphones, as well as a screen capture program for recording gameplay).

5. The game begins, with the director verbally guiding the matcher through the virtual atoll environment until all treasure chests have been discovered, or until the researcher or participants decide to stop (for example, if the game has become too time-consuming).

6. If time permits, the players may swap roles and steps 1-5 may be repeated with a second game scene. This scene has identical geography to the first, but the treasure chests are hidden in different locations (accordingly, there is a different demonstration video for this iteration).
3.2 Appearance and gameplay

In order to maximise immersion in the virtual environment, the VAT employs a first person perspective (as opposed to, for example, the third person view or the ‘over the shoulder’ perspective found in many video games); therefore, the players’ experience of the virtual environment is as similar as possible to the way they experience their real environment. The scenes in the VAT were created by an ‘Island Generator’ program designed using the Unity platform for computer game design. With minimal training, it is possible to use the Island Generator to create and terraform one’s own virtual atoll and/or island environments.

The environment generated for the VAT consists of an atoll at some distance from a large non-atoll island. The latter was included so that players do not travel only within the confines of the lagoon, but have the chance to leave the atoll and

---

8 Unity is available for free download from www.unity.com and is required to run the Island Generator program.
engage in long-distance ‘ocean-voyaging’, a practice which in some languages requires different navigational strategies (see Palmer, 2007 for Marshallese; François, 2004 for Oceanic in general). In order to help the players differentiate between the lagoon and the ocean, the game’s ocean water is a deeper blue than the lagoon water, the ocean side of islands is considerably rockier, and piers – of which there are three around the virtual atoll – are always on the lagoon side of islands. These features parallel the geography of real-world atolls. While many of the game’s islands are similar in appearance to one another, none are actually identical, and a number of features including vegetation, rocks, buildings, vehicles, shape, size and location may be used to tell them apart. The game also features a prominent sunset to allow participants to derive cardinal directions from the position of the sun, should this occur to them.9

Across both scenes, the treasure chests are hidden inside houses, on balconies, underneath piers, in the back of a van, or on the ground (either concealed by other objects or out in the open). The chests were deliberately placed in a wide variety of locations in order to facilitate participants’ memory of them. The game includes a counter at the top left of screen which indicates how many chests are yet to be opened.

The controls for the VAT were designed to be as simple as possible, under the assumption that most players would be unfamiliar with computer games, if not computers in general. Therefore, while most modern games use either the arrow keys or the ‘wasd’ keys for movement, the VAT only requires one keyboard key (the space bar), which moves players forward in the direction they are currently facing. The view is controlled by moving the mouse, and objects (i.e., treasure chests and speedboats) are engaged by pressing the left mouse button. In addition to the relative simplicity of having the mouse control orientation, it is also advantageous in that the mouse is presumably less likely than the arrow keys to bias participants towards using egocentric spatial terminology such as ‘left’ and ‘right’. A disadvantage, however, is that players are unable to move backwards, and instead must rotate the view 180 degrees before again pressing the space bar.

9 If desired, the sunset sky can be switched to a daytime one in the Island Generator program.
4. Results and discussion

4.1 The data

To date, we have employed the VAT as an elicitation game with 12 Dhivehi-speaking pairs from Laamu (Haddummati) Atoll, Maldives, and seven Marshallese-speaking pairs from Jabor, Jaluit Atoll, Marshall Islands. In Laamu Atoll it was possible to administer the VAT in six different islands, though most pairs were recorded in the capital island of Fonadhoo (five pairs) or the fishing/farming island of Dhan’bidhoo (three pairs). Owing to the need to recruit participants who had some basic familiarity with computers, most participants from both language communities were young men. A version of the task with five treasure chests was used for most pairs, though a newer version with only four chests was used for four of the Dhivehi pairs in order to reduce the duration of the task.\textsuperscript{10} The combined Dhivehi data is approximately six hours in total, of which a sample of one hour and 11 minutes was transcribed and translated for analysis. The Marshallese data amounts to approximately two hours and 10 minutes, all of which is fully transcribed and translated. In addition, a second field trip to the Marshall Islands is planned for 2015/2016 during which more VAT data may be collected.

4.2 Linguistic results

On the basis of the data collected, it appears that the VAT was partly successful in eliciting the kinds of spatial descriptions for which it was designed, though the nature of the task appeared to favour certain spatial strategies over others. For both languages, the VAT succeeded in eliciting some spatial vocabulary that was not observed in small scale spatial tasks such as the Man and Tree game. For example, the distal locative demonstrative \textit{ijuweo} ‘over there, yonder’ was

\textsuperscript{10} The newer version also has a number of small improvements, including a more obvious contrast between the lagoon and ocean, and smoother interaction with in-game objects.
produced by several Marshallese speakers, as were topographically-based motion verbs such as *buňlik* ‘sail out into the ocean’. In Dhivehi, the motion verbs *arān* ‘ascend’, *erēn* and *faibān* ‘descend’ were commonly used to describe getting on and off islands, even when the participants were well inland and the ground was flat. The use of such terms suggests that participants were indeed identifying the virtual atoll environment with real atoll environments, which had been an important aim in designing the task. Further evidence of this is in some of the jokes and comments made by participants during the task, explicitly identifying virtual islands with real-world ones, as in the following example from Marshallese:


Face sun go.straight 1PL.INCL -two go to Majuro

‘Face the sun. Go straight. We are going to Majuro.’

*e pomi Jaluit, ƚe? E- nañin to an jāde? … Eĩñe!*

Where Jaluit man 3SG almost long.time 3SG.POSS appear there.it.is

Where is Jaluit, man? How long until its appearance? … There it is!’ 11

In addition, the Marshallese data has a few instances of players employing the directionals *iar* ‘lagoonward’ and *lik* ‘oceanward’, though it is not always clear if the terms refer to an axis in the virtual environment or to the lagoonward-oceanward axis in the external world. Despite our best efforts to make the virtual lagoon and ocean as visually distinct as possible, in some cases participants felt

---

11 Majuro is the capital of the Marshall Islands, an atoll some 200km away from Jaluit Atoll, the participants’ home atoll where the task was recorded.
unable to extrapolate a lagoonward-oceanward axis within the game, as the following exchange demonstrates:

(2) YM1

\[ e- \ pād \ ilo \ \textit{jekā} \ \textit{kañe}. \]

3SG- be.located at some.place those

‘It is somewhere around there.’

TP1

\[ \textit{lik} \ \textit{ke} \ \textit{iar} ? \]

ocean.side or lagoon.side

‘Ocean side or lagoon side?’

YM1

\[ [...], \textit{je-ŋak} \ \textit{ewi} \ \textit{iar} \ \textit{im} \ \textit{lik}. \]

bwe

so 1PL.INCL-not.know where lagoon.side and ocean.side

[...] ‘So we don’t know where the lagoon side and the ocean side are.’

MAR_VAT_JAB_20141129_YM1_TP1_W: 06.52

Perhaps due to perceived ambiguities in using topographically-based directionals for such a task, the participants often deferred to other strategies. The Marshallese participants primarily relied on left-right terms, real-world cardinal directions, or somewhat vague uses of locative demonstratives. Dhivehi speakers also relied very heavily on left-right terms and demonstratives, and many descriptions lacked frame of reference information altogether, as in the following example:

(3) \[ \textit{goho-bala} \ \textit{ēnū} \ \textit{fālam-ek} \ \textit{lā-fa} \ \textit{raš-ek} \]

\[ \text{go-IMP sit.there.PST.FOC pier-INDF put-PTCP island-INDF} \]

‘Go, there is an island with a pier.’

DIV_VAT_LF_20140129_1_1_AZ2_IS3_SW: 04:47
However, a smaller number of Dhivehi descriptions invoked topographical or cardinal directions. Topographical terms included *eggamu* ‘inland’, *atiri* ‘beach’, and *fuṭṭaru* ‘reef’; these were used exclusively to refer to features of the virtual environment and to directions within that environment. Interestingly, there was no clear evidence of a lagoonward-oceanward axis in the Dhivehi data, in contrast to Marshallese where such a directional system was represented, even if it was not a dominant strategy. Dhivehi cardinal terms were occasionally used for virtual cardinal axes (typically ‘east’ or ‘west’ when facing the setting sun), but more often for real-world directions, with ‘north’ corresponding to the screen’s left if the players were facing east, for example.

There are a number of possible explanations for the preference in both samples for egocentric strategies such as left-right terms and demonstratives. Firstly, the participants recruited were predominantly young adults who had some level of computer literacy, and other experiments such as the Man and Tree game revealed that this group was more likely to use egocentric spatial terms in other tasks too. But one might also suppose that the very use of a computer encourages egocentric frames, because even if the controls and interface for the VAT do not have left-right asymmetries, participants may nonetheless associate computers with left-right systems. This is because other computer-based activities require the use of such systems (e.g., text directionality, left vs. right clicking, arrow keys, etc.). Finally, although participants were familiar with the type of environment represented in the VAT, they were unfamiliar with the geographic details of the atoll in the game (e.g., which islands contained houses, etc., as well as the locations of the islands themselves). Since participants could only have developed a fairly rudimentary and inaccurate mental map of the virtual atoll from their exposure to it in the brief training phase, it would have been challenging for them to relate their current position in the game to other in-game locations that were out of view. This may have fostered strategies that focused on the participants’ immediate movements, without any reliance on more distant features such as the virtual lagoon or sun. It would be interesting to observe how participants play a version of the game after they have acquired much more experience and
familiarity with the geography of the virtual atoll, and this may be a future application of the task.

4.3 The VAT as an elicitation tool: strengths and weaknesses

When trialling the VAT, several limitations emerged. Firstly, despite our efforts to make the controls as simple as possible, many participants still found them challenging, particularly in parts of the game which required finer control, such as climbing up the staircases of houses. These were generally participants who did not have much experience operating computers. Because of this, we mostly recruited participants who were young and sometimes more highly educated than other members of the community. Thus the data elicited was not necessarily representative of the wider speech community, but only of a subgroup, and it may fail to capture more traditional patterns of usage as favoured by older speakers of Dhivehi and Marshallese. For some languages or at some field sites, presumably, it would not be possible to recruit any suitable participants at all, if nobody in the community has a sufficient level of computer literacy.

A further issue was the duration of the task, which was often over an hour (including instructions, training, and two games). While this did not cause too many problems at our field sites, it could potentially be problematic if participants are busy or have limited attention spans. It would be possible to mitigate this by running the task across two sessions, with the participants swapping roles for the second session.

Finally, when participants got lost in the virtual world or forgot the location of the treasure chests, the game sometimes descended into a relatively quiet and/or uninteresting phase, with participants meandering around silently as they attempted to find the chests by trial and error. As mentioned earlier, this can be remedied to an extent by reducing the number of chests to four, but occasionally one may nonetheless finish with a paucity of meaningful data, at least for some parts of the game. In contrast, the design of previous tabletop space games precludes the use of a trial-and-error strategy, since the split field of view prevents
players from knowing if they have made a correct ‘match’ until the end of the task.

On the other hand, the VAT has a number of advantages over earlier spatial language elicitation tasks, avoiding many of the shortcomings outlined in §2.2. In addition to the task’s ability to elicit larger scale spatial descriptions (as demonstrated in the previous section), the VAT is in many ways a more realistic task in comparison with other space games, despite the fact that it takes place on a computer screen. This is partly because of the first-person perspective experienced by participants in the game, which is the perspective one would ordinarily have when giving or receiving spatial descriptions in real life. Previous space games involving a bird’s-eye view of small objects within reach provide a somewhat less naturalistic context for precise spatial language. Such games are also artificial in that they divide participants with a screen, bringing about a split field of view. Although this is fundamental to making certain games succeed as elicitation tasks, a divided field of view of identical objects is highly unusual in real life, and can easily lead to misunderstandings that would never occur in normal conversation (for example, in the Man and Tree game, if a director says the equivalent of ‘The toy man is facing me’, the matcher will often be confused as to whether this means towards the screen – i.e., perpendicular to the participants – or towards herself).

5. Conclusions

We submit that the VAT is a valuable tool for eliciting spatial language, and in particular for testing Palmer’s (2015) Topographic Correspondence Hypothesis. At our field sites in the Maldives and the Marshall Islands, it successfully elicited certain large scale spatial descriptions that were not prompted by more conventional tasks, and provided meaningful data that will go some way to evaluating the Topographic Correspondence Hypothesis, as well as to answering more general research questions on the relationship between spatial language and the environment (though we cannot hope to resolve those questions here).
The VAT overcomes many of the shortcomings of more established spatial language elicitation tasks, though it has certain limitations of its own. However, many of these can be mitigated, or will become less problematic over time (e.g., in the case of difficulties with operating the controls, it is likely that the pool of potential participants for the task will grow as computer literacy continues to spread into even the remotest of communities). In any case, we regard the VAT as a useful addition to other space games, rather than as a replacement. Table 1 below summarises the ways in which the VAT differs from previous tabletop space games:

Table 1: Comparison of the VAT to previous spatial language elicitation tasks

<table>
<thead>
<tr>
<th></th>
<th>Previous space games</th>
<th>VAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>Small, tabletop</td>
<td>Medium-large (simulated)</td>
</tr>
<tr>
<td>Field of view</td>
<td>Divided</td>
<td>Shared</td>
</tr>
<tr>
<td>Perspective</td>
<td>Bird's-eye view</td>
<td>First person</td>
</tr>
<tr>
<td>Stimulus materials</td>
<td>Artificial, abstract</td>
<td>Less artificial or abstract</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Objects easy to manipulate</td>
<td>Controls somewhat difficult</td>
</tr>
<tr>
<td>Specificity of descriptions</td>
<td>Highly specific</td>
<td>Less specific (trial-and-error strategy possible)</td>
</tr>
<tr>
<td>Source of director's knowledge</td>
<td>Separate but identical stimuli</td>
<td>Memory</td>
</tr>
</tbody>
</table>

Finally, although the VAT is currently limited to atoll and island environments, it would be possible to use Unity or other software to build alternative virtual worlds. It would be intriguing, for example, to trial a riverine version with speakers of a language spoken in that environment, or an alpine map in a mountain-based community. Additionally, the VAT, along with the Island Generator program used to create the virtual environment, has the potential to be
adapted to the needs of the researcher and could foreseeably be used for a variety of purposes, such as studying the development of cooperative communication strategies,\textsuperscript{12} or for more general linguistic elicitation. Since relatively few researchers are likely to take the VAT to atoll-based communities, we anticipate that the immediate future of the task lies in these other applications of it, as well as in non-atoll island communities for which the Island Generator program can already build suitable virtual environments.

\section*{Acknowledgements}
Funding for this project was provided by the ARC DP grant G1100293: ‘Thinking and talking about atolls: the role of environment in shaping language and our understanding of physical space’. We are grateful to Bill Palmer, Alice Gaby and one anonymous reviewer for many helpful suggestions, and to all the Maldivian and Marshallese consultants who agreed to participate in this project. We must also acknowledge one other individual – who has chosen to remain anonymous – for contributing the technical expertise and ongoing support to the Island Generator software that made the Virtual Atoll Task possible.

\section*{References}


\textsuperscript{12} We thank Mark Ellison for this suggestion.


Epistemic modulations and speakers stance in Cantonese conversations

Winnie Chor
The Open University of Hong Kong
wowchor@ouhk.edu.hk

Abstract. Based on data obtained from 20 interview recordings that involve descriptions of pictures of different attractions, this paper examines how native speakers of Cantonese use different strategies, in particular grammatical strategies, to modulate (i.e. upgrade or downgrade) the epistemic strength of their claims when they are challenged or disaffirmed by their addressee(s) in an interactional context. Our analysis reveals that speakers make use of four categories of strategies to modulate their assertions, including i) epistemic modals (e.g. jing1 goi1 ‘should’, bo2 nang4 ‘might’), and adverbials (e.g. kei4 sat6 ‘in fact’, si6 sat6 soeng6 ‘in fact’, tung1 soeng4 ‘usually’), ii) evidentials which are explicit (e.g. ngo5 gu2 ‘I guess’) or implicit (e.g. hou2 ci3 ‘seems’, jau5 bo2 nang4 ‘likely’), iii) sentence final particles with weak epistemic strength (e.g. ge2, gwaa3, za1 maa3) or strong epistemic strength (e.g. wo3, gaa3), iv) discourse-pragmatic strategies (e.g. to keep silent, to provide a reason, as well as the use of redressive language).

Keywords: evidentiality, epistemicity, stance-marking, Cantonese

1. Introduction
All natural languages have means to convey meanings of various kinds, from those which are more objective or impersonal (e.g. meanings concerning time and quantity), to those which are more subjective, conveying different shades of speaker’s moods and perspectives (e.g. how the speaker evaluates a certain
situation, how evident the speaker’s conclusion of a situation is, etc.). Knowing how to interpret the speaker’s subjective stance and attitude is vital in everyday communication. When we interact with other people, we need to understand not only the word meaning, but also the speaker meaning – we constantly need to recognize the speaker’s intention and attitude. There is now a growing body of literature on stance-taking phenomenon from various perspectives, typically from the sociolinguistic and anthropological perspective (Clayman & Heritage 2002; Fetzer & Fischer 2006; Englebretson 2007), as well as from the discourse-pragmatic perspectives (Mori 1999; Wu 2004; Yap & Lun 2010; Iwasaki & Yap 2015). Findings from these studies have important contributions to our understanding of the kinds of strategies that speakers employ when they need to convey their thoughts and attitudes to others.

Building on the more general works, earlier studies on stance-taking have been predominantly on evidentiality marking, understood as a pragmatic and grammatical domain that specifies the ‘source of evidence’ for a given event, whether it was seen, directly heard, told, or experienced (Willett 1988; Aikhenvald 2004). Studies on evidentiality have been done from different perspectives. For example, Woodbury (1986), Izvorski (1997) and Lee (2013) studied the connection between temporal elements and evidentiality. Willett (1988) and Malone (1988) revealed the origins and the development of evidentiality. Nuyts (2001) and Cornillie (2007) investigated the role of (inter)subjectivity in epistemic expressions, and analysed its relationship with evidentiality. Mortelmans (2001), Cornillie (2005) and Hara (2006) on the other hand, focused on the evidential dimension of modals in German, Spanish and Japanese respectively. Aikhenvald & Dixon (2003) and Aikhenvald (2004) focused on the typology of evidentials, justifying that evidentiality is in fact a category not overlapping with other categories, such as modality.

In more recent years, the scope of evidentiality studies has been expanded to see what other functions evidential markers can serve besides ‘source of evidence’, which goes beyond our original understanding of evidentiality. A number of studies on evidentiality have been undertaken in Asian languages such as Chinese (including Cantonese and Mandarin), Japanese, and Korean from a discourse
perspective, and a majority of these studies have paid special attention to ‘say’-derived evidentials. As far as Chinese languages are concerned, Wang et al. (2003) on Mandarin shuo ‘say’ have investigated into how shuo as a say verb can be used as a counter-expectation marker when appearing utterance-finally, while Leung (2010) has provided a very detailed account of how the various functions of the Cantonese waa6 ‘say’ have come about diachronically. With regard to Japanese, S. Suzuki (1999) and R. Suzuki (2007), for instance, have identified that tte as a quotative marker has been reanalysed as a pragmatic marker when occupying a sentence-final position, expressing a wide range of speaker attitudes. Works by Kim (2005, 2011) on Korean, on the other hand, have explored how evidential markers can be used to express epistemic meanings in a discourse context. It is found that hearsay evidential markers such as tamye and tako can be used by speakers to express their epistemic stance. Particularly, the markers are used to negotiate the speaker’s epistemic footing in an interactional discourse.

All these previous studies have paved the way for us to expand on the research on evidential marking, and to go beyond evidentiality alone to see how it is relevant to epistemicity and speaker attitude in stance-marking. While some recent studies have shown that epistemic modality and evidentiality are closely related, with evidential markers often used in conversations as discourse-pragmatic markers to modulate the strength of the speaker’s epistemic claim and to help externalise his or her attitude (Kim 2005, 2011), and other studies on evidentiality have explored how participants of a conversation can convey neutrality, engagement, or detachment towards an utterance (Fox 2001; Kim 2005, 2011; Cornillie 2009; Tang 2010; inter alia), studies looking at how the evidential system interacts with other grammatical systems to generate speaker-oriented meanings are still scarce. To fill this gap, the present study aims at investigating the wide range of strategies that speakers of Cantonese deploy to display their epistemic stance and subjective attitude in daily conversations. Based on data obtained from contemporary corpora composed of natural conversations and interviews, we attempt to see how various grammatical resources and strategies can be used to externalise the speaker’s subjective mood as well as how the epistemic commitment is modulated, from a discourse-pragmatic perspective.
This paper is organised as follows. Section 2 describes the data and the methodology for this study. Section 3 provides a brief discussion of stance-marking as a pragmatic phenomenon and how it is understood in relation to the domains of epistemicity, evidentiality and attitude. Having a good grasp of these concepts will facilitate the understanding of how an interlocutor’s epistemic strength can be reinforced or mitigated. Section 4 discusses how epistemic modulations are done in Cantonese conversations and specifies the concrete strategies for epistemic upgrading and downgrading. Section 5 summarises the findings of the investigation, and concludes with a brief discussion of the implications of the present study for the field of stance-marking.

2. Data and methodology

The data for this study consist of recordings of interviews between native Cantonese speakers. The conversational data come mainly from two projects, namely, *Epistemic Modulation and Speaker Attitude in Cantonese: A Discourse-Pragmatic Perspective* and *Stance-Marking in Asian Languages: Linguistic and Cultural Perspectives*. A ‘picture kit’ has been designed to elicit data in both projects as part of their data collection. Each ‘picture kit’ comprises a set of pictures depicting different buildings and attractions, such as mosques, European-style castles, Asian-style temples, and different towers of the world. These pictures are designed to elicit a range of expressions involving the speaker’s stance in the context of an individual interview. Altogether 20 interviews (about 30 minutes each) were arranged. Ten of the interviews involved students from the Open University of Hong Kong (OUHK) and ten involved students from the Hong Kong Polytechnic University (HKPU). In the interviews, participants were asked to talk about where they

---

1 *Epistemic Modulation and Speaker Attitude in Cantonese: A Discourse-Pragmatic Perspective* is a project funded by the Research Grant Council of Hong Kong (FDS #UGC/FDS16/H07/14), awarded to the author of this paper. The project aims to explore the various grammatical resources and strategies which Cantonese speakers use to externalise their subjective mood and to modulate their epistemic commitments in an interactional context. *Stance Marking in Asian Languages: Linguistic and Cultural Perspectives* is another project also funded by the Research Grant Council of Hong Kong (GRF #PolyU 551310; PI: Dr Foong-ha Yap). This is a cross-linguistic and cross-cultural study attempting to uncover the common tendencies in the use of stance-marking strategies across Asian languages. The corpus can be accessed at this website: http://www.engl.polyu.edu.hk/research/corpus/corpus.htm.
thought the pictures were taken, and to support their claims with reasons. Since buildings such as castles and temples are not uniquely associated with a particular place, participants thus needed to make a guess. For instance, mosques in Indonesia, Pakistan and Bangladesh might look alike. When a picture of a mosque was shown to the participants, they would have the opportunity to talk about the picture, and to express their subjective viewpoints about its location. They would need to support their claims with reason, and to indicate how committed they were to their claims. When they were challenged by the interviewer, they might need to either upgrade or downgrade their assertions. In this way, strategies that speakers use to modulate their epistemic claims can then be identified.

3. Expressing Stance
In our daily conversations, we provide many linguistic cues for others to interpret our stance. The stance of a speaker is understood as the epistemic or attitudinal comment on propositional information (Biber 2004). In other words, it deals with how the speaker conveys his or her value judgments, personal feelings and degree of commitment to the truth value of a given proposition (Englebretson 2007). A speaker’s stance includes subjective expressions of his or her mood, attitude, assessment, and perspective, and is reflected at various levels, including lexical, phrasal, and clausal levels (Yap and Lun 2010). Speakers can display their stance through the use of many different strategies, such as the use of different lexical/phrasal choices, syntactic structures, pragmatic markers, intonation/prosody, as well as various discourse strategies including hedging devices and even silences. For example, in English, speakers deploy a rich array of expressions to indicate their epistemic stance. These epistemic strategies include the use of adverbials (e.g. probably, certainly), modals (e.g. may, might), epistemic

---

2 The participants are all university undergraduate students aged 18-23, growing up in Hong Kong speaking native Cantonese. Two interviewers were involved in the Epistemic Modulation project and two in the Stance Marking project. They were all research assistants of the project, also speaking Hong Kong Cantonese as their mother language. The interviewer and the interviewee did not know each other in person. All interviewees were given a consent form to sign before the interview so they knew that the conversations recorded will be used for research purposes. However, they were not told as to how the data was going to be used, so they did not know that they were going to be deliberately challenged.
phrases (e.g. *I think*, *I believe*), as well as other lexically transparent expressions (e.g. *I really love that*) and grammaticalised expressions (e.g. *I’m afraid*), to explicitly display their epistemic attitude and subjective mood.

In addition to epistemic modality, speakers also often indicate their degrees of commitment to their claims by acknowledging the source and reliability of their information, which can be lexically encoded, grammatically encoded, or heard, or simply inferred. Speakers may try to be explicit by saying *As stated in the most recent report*, or, they might choose to be more vague and thus deploy impersonal evidential expressions such as *People say* or *It is said*. As Aikhenvald remarks, ‘[i]n about a quarter of the world’s languages, every statement must specify the type of source on which it is based – for example, whether the speaker saw it, or heard it, or inferred it from indirect evidence, or learnt it from someone else.’ (Aikhenvald 2004:1). Cusco Quechua (Faller 2006), St’át’ímcets (Matthewson, Henry & Hotze 2007) and German (Diewarld & Elena 2010), to name a few, are the languages in which the highly developed and finely calibrated evidentiality systems are observed.

We have seen that stance-marking has established itself as a research area and has attracted a wealth of research studies carried out from a wide range of different perspectives. Works by Rhee (2011, 2012) on Korean have shown how stance-marking is relevant to the nominalisation process and different sentence types. S. Suzuki (1999) and R. Suzuki (2007), on the other hand, have identified how the Japanese quotative marker *tte* has been reanalysed as a sentence-final particle to express a wide range of speaker attitudes. Studies by Chor (2011, 2012) on Cantonese explore how stance-marking can be achieved by the use of perception verbs and verbal particles. Other recent studies have also shown that epistemic modality and evidentiality are closely related when expressing the speaker’s stance. Ekberg and Paradis (2009), for example, is an edited volume that explores the nature of evidentiality and the relation of evidentiality to epistemic modality, as well as the mappings of linguistic expressions of evidentiality to meaning, communicative function and reasoning. Works by Kim (2005, 2011) on Korean have also explored how evidentiality is relevant to epistemic strategies in a discourse context.
Inspired by Kim’s works on evidentiality, especially her 2011 work on negotiating epistemic rights in Korean, the present study will look into how speakers modulate their epistemic claims from a discourse-pragmatic perspective. While Kim (2011) has focused on how the evidential marker -*tamye is used as a discourse-pragmatic marker to manage epistemic rights of the participants, our study will focus more on how and for what reasons speakers might also use other grammatical strategies, such as grammaticalised epistemic/evidential expressions, particles, as well as various discourse strategies, to upgrade/downgrade their assertions when their claims are queried or challenged.

4. Negotiation of epistemic strength
A speaker’s epistemic stance is being negotiated and monitored continuously in an interactional context. Based on the sets of interview data in Cantonese, a range of epistemic modulation strategies have been revealed. We believe that interview data of this kind is particularly useful in the context of our present study. A power asymmetry is often involved in such context – the interviewer has more information and thus has more control over the interview, while the interviewee has less information and thus is less powerful in the interview. When the interviewee’s claims are queried or challenged, he or she will need to give explanations and to provide evidence to support the claims. By doing this, a lot of stance-marking strategies would have been elicited. Also, the context of an interview will allow the interviewees to express their views freely, resulting in the conversational data collected being natural and unimpeded. The present study explores how these different types of strategies can collaborate with one another to calibrate the speaker’s epistemic claims in natural, interactional contexts.

4.1 Strategies of epistemic modulations
Previous studies on epistemic stance have been predominantly done in relation to the evidential system. Without doubt, evidential markers play a very important role in negotiating the speaker’s stance. However, other strategies are equally important and cannot be neglected. Altogether four categories of strategies were
identified in our data. They include: i) epistemic modals (e.g. jing1goi1 ‘should’, bo2nan4 ‘might’) and adverbials (e.g. kei4sat6 ‘in fact’, si6sat6soeng6 ‘in fact’, tung1soeng4 ‘usually’), ii) evidentials which are explicit (e.g. ngo5 gu2 ‘I guess’) or implicit (e.g. bou2ci3 ‘seems’, jan5 bo2nan4 ‘likely’), iii) sentence final particles with weak epistemic strength (e.g. ge2, gwaa3, za1maa3) or strong epistemic strength (e.g. wo3, gaa3), iv) discourse-pragmatic strategies (e.g. to keep silent, to provide a reason, to be indirect, as well as the use of hedging devices and redressive language).

4.1.1 Epistemic modals and adverbials

In English, there are various ways that we can express a certain attitude towards a proposition, including the use of modal auxiliary verbs (e.g. may, might), modal adverbs and adjectives (e.g. likely, possibly, probable), as well as verbs that signal a varying degree of speakers’ commitment (e.g. know, think, doubt). These complement taking verbs have grammaticalised into epistemic adverbials when they are used with the first person pronoun I. Examples include the English I think, which could be used in the epistemic sense of ‘probably’ in flexible syntactic positions, including the utterance-initial, utterance-final and utterance-medial positions (Thompson & Mulac 1991; Karkkäinen 2003).

As observed from our data, Cantonese speakers also frequently make use of modal auxiliary verbs, as well as adjective and adverbs, to externalise their epistemic commitment to the information conveyed in their utterance.

(1)

01 IR: Dim2gaai2 nei5  gok3dak1 hai6  Zung1Gwok3sin1?

Why  2SG feel  COP China  SFP

3 The use of I think is also closely related to the expression of inferential evidential meaning. In fact, the semantic domains of evidentiality and modality are closely related to each other. This will be explored in section 4.1.2.
‘Why do you think it is China?’

02 IE : 

\[\text{Jan1wai6... jan1wai6 (explanation omitted) ngo5nam2jing1goi1 don1 hai6}\]

Because because (explanation omitted) 1SG think MOD also COP

\[\text{wui6 hai6 Zung1Gwok3 jap6-min6go2 di1 gu2zik1 gam2 joeng2 laa1.}\]

MOD COP China inside that CL monument such look SFP

‘Because...because (explanation omitted) I think it is possible that it would be some monument in China.’

In (1), the interviewee is being challenged as to why she thinks the temple in the picture is from China. The speaker deploys the psych verb nam2 ‘think’, the modal auxiliary jing1goi1 ‘should’ and wui5 ‘would’, and in collaboration these strategies function to downgrade the speaker’s epistemic commitment to the proposal that the temple is from China. Instead of saying nei1 go3 hai6 zung1gwok3 jap6-min6 go2di1 gu2zik1 ‘this one is a historic spot inside China’, the speaker expresses that she only thinks (i.e. not sure) that the temple should be (i.e. probably but not absolutely) a historic spot in China. Thus, emerging in here is only a weak epistemic interpretation.

Worth noting is that, as a lexical verb, gok3dak1 ‘feel’ can be used as a psych verb to introduce a complement clause that expresses one’s feeling or thought. It can also be used to convey the speaker’s psychological condition/state of mind. In an interactional context, the use of ngo5 gok3dak1 ‘I feel’ has extended to the cognitive domain, and is best interpreted in the epistemic and inferential sense similar to the use of ‘I think’ or ‘I guess’ in English.

---

4 The examples cited in this paper include morpheme-by-morpheme gloss and English translations of the utterances. The abbreviations used include: CL = classifier; CONN = connective; COP = copula; EMPH = emphatic marker; EXIST = existential verb; EXP = experiential; GEN = genitive; INJ = interjection; MOD = modal; NEG = negator; PRT = particle; Q = interrogative marker; RVC = verbal resultative complement; SFP = sentence final particle; SG = singular; 1 = first person; 2 = second person; 3 = third person. The Romanisation scheme that we use here is based on the Linguistic Society of Hong Kong Cantonese Romanisation System.
(2)

01 IR: Dim2gaai2 gok3dak1 hai6 Jat6Bun2 ne1? Dim2gaai2m4-hai6

Why feel COP Japan Q Why NEG-COP

Zung1Gwok3 m4-hai6 Hon4Gwok3 aa3?

China NEG-COP Korea Q

‘Why do you think it is Japan? Why isn’t it China or Korea?’

02 IE: Ngo5 gok3dak1 gam2gok3-soeng6 hou2ci3 m4-ci3.

1SG feel feeling-up seem NEG-resemble

‘I think that by intuition it seems not [to be China or Korea].’

In (2), ngo5 gok3dak1 is an epistemic phrase that the speaker uses to indicate his stance of probability and uncertainty. Like the epistemic phrase I think in English, ngo5 gok3dak1 also enjoys the freedom of syntactic flexibility. The phrase ngo5 gok3dak1 can be inserted in the below three positions as indicated by the blanks:

__gam2gok3 soeng6 __hou2ci3 m4ci3__

4.1.2 Evidential devices

The domains of epistemic modality and evidentiality are closely related to each other. From the diachronic perspective, epistemic modality often develops into the inferential domain, and it is the inferential uses of epistemic modality that intersect with evidentiality (i.e. the speaker’s inference can be treated as a kind of evidence supporting the speaker’s claim). The involvement of speaker’s inferences as the source of information is also an important aspect of the speaker’s stance. We might expect that claims with a weaker epistemic strength is also less subjective and relies more on inferential/indirect evidence, while claims with a stronger epistemic strength would be reinforced by direct evidentials and markers
of higher levels of subjectivity. It is found in our data that markers of evidentiality when used in an interactional context can claim or mitigate the speaker’s epistemic commitment over the information, in response to the queries or challenges made by the other participant. Consider example (3) below.

(3)

01 IR: Ni1 zoeng1 Baa1Lai4Tit3Taap3 aa3?
   This CL the-Eiffel-Tower Q
‘Is this one the Eiffel Tower?’

02 IE: Hou2ci5 m4-gok3 Baa1Lai4Tit3Taap3 hai6 gam2 ge3 ngaan4sik1,
   Seem NEG-feel the-Eiffel-Tower COP such GEN colour
   jan3zoeng6-zung1 keoi5 hou2ci5 daai6koi3 hai6 baak6sik1 aa3…
   impression-in 3SG seem approximately COP white SFP
‘It seems like I do not think that the colour of the Eiffel Tower is like this. In my impression, it seems, approximately, it is white…’

In (3) above, together with the adverbial phrases m4 gok3 (I)-neg-feel ‘(I) don’t think (that)’ and jan3zoeng6zung1 ‘in my impression’, the speaker utilises the implicit evidential hou2ci3 ‘seems’ as a mean to downgrade his comment that the picture might not be the Eiffel Tower. By using the evidential marker, the speaker has downgraded his degree of epistemic commitment to his comment regarding the colour of the Tower.

4.1.3 Sentence final particles

Compared to northern varieties such as Standard Mandarin, Cantonese is known to have a much wider range of particles, in particular sentence final particles (SFP). Sentence final particles are the bound morphemes optionally attached to
the end of a sentence (Law 1990; Matthews & Yip 1994; Cheung 2007) or an utterance final position (Luke 1990; Fung 2000; Leung 2005). They are responsible for different syntactic, semantic and pragmatic functions. When they are used, most often they can form particle clusters (Yau 1980; Law 1990; Matthews & Yip 1994; Cheung 2007), combining up to seven single particles (Leung 2005: 1-2). Substantial works have been done on sentence final particles. Some studies have focused on the core semantic meanings of these particles based on corpus (Leung 2005) and phonological properties (Fung 2000). Some have concentrated on a particular (group of) particle(s), e.g. the focus particles $zaa^3$ and $tim^1$ (Law 2004), the quotative particles $wo^3$, $wo^5$ and $bo^3$ (Leung 2010), the particle $sin^1$ (Tang 2012), just to name a few. Other studies might have taken a more formal approach to classify particles into categories according to their syntactic functions (Law 1990; Law 2002; Sybesma & Li 2007).

Sentence final particles form a very powerful grammatical system in Cantonese. They are often employed in conversations to express a wide range of meanings in the areas of aspect, modality, evidentiality, as well as speaker stance, among others. The speaker’s subtle intention can be accomplished by the skilful selection of just one particle. Consider how, as seen in Table 1 below, the addition of different sentence final particles can affect the meaning of the utterance $Keoi^5 wui^5 heoi^3$, literally ‘s/he will go’.

**Table 1. Cantonese SFPs with their meanings translated in English**

<table>
<thead>
<tr>
<th>Base utterance</th>
<th>Sentence final particle (SFP)</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Keoi^5 wui^5 heoi^3$</td>
<td>$wo^5$</td>
<td>‘People say that he will go.’</td>
</tr>
<tr>
<td></td>
<td>$gwaa^3$</td>
<td>‘It is probable that he will go.’</td>
</tr>
<tr>
<td></td>
<td>$ge^3$</td>
<td>‘It is certain that he will go.’</td>
</tr>
<tr>
<td></td>
<td>$gaa^3$</td>
<td>‘For sure that he will go.’</td>
</tr>
<tr>
<td></td>
<td>$me^1$</td>
<td>‘He will go? I doubt it.’</td>
</tr>
</tbody>
</table>

The particles listed in this table are by no means exclusive. Cantonese has a very rich inventory of sentence final particles. Previous studies have identified a large number of them, ranging from over 25 to as many as 95 (depending on how these particles are defined). For extended readings on Cantonese sentence final particles, please refer to previous works such as Yau (1980), Kwok (1984), Law (1990), Luke (1990), Fung (2000), and Leung (2005).
From the above table, we can see how speakers of Cantonese use sentence final particles to indicate subtle nuances of meaning. Some particles signal varying degrees of speaker’s commitment. For example, gaa3 and ge3 indicate high epistemic certainty, while gwaa3 indicates uncertainty. Wo5 is a hearsay particle indicating that the source of information is from elsewhere, i.e. from someone other than the speaker. As far as ‘yes/no’ interrogative particles, bo3 expects a positive answer whereas me1 expects a negative answer.

As observed from our data, particles like lo1, gaa3, or laa1 are often used to upgrade the speaker’s epistemic strength, whereas particles like ge2 or gwaa3 have weaker epistemic strength and are often used to mitigate the speaker’s epistemic commitment to the information reported. See examples (4) and (5) below.

| bo3 | ‘He will go, right?’ |

(4)

01 IR: Ni1 zoeng1 jat1 ding6 m4-bai6 Zung1 Gwok3?

This CL definitely NEG-COP China

‘Are you sure this one is not China?’

02 IE: Jat1 ding6 m4-bai6!

Definitely NEG-COP

‘I’m sure it is not!’

03 IR: Gam3 hang2 ding6?

That sure

‘Why are you so sure?’

04 IE: Keoi5 m4-wui6 zing2-dou2 ni1 di1 gam2 ge3 fung1 gaak3 gaa3 wo3.

3SG NEG-MOD make-RVC this CL such GEN style SFP SFP

---

6See Leung (2010) for the grammaticalisation of wo5.
‘They [the Chinese] are not able to build with such style.’

(5)

01 IR: Gam2 ni1dou6 bin1dou6 aa3?

Then here where Q?

‘Then, where is it?’

02 IE: Ni1 go3 hou2 ming4bin2 laa1, zau6 hai6 Zung1Gwok3 ge3

This CL very obvious SFP EMPH COP China GEN

Maan6Lei5Coeng4Sing4 gwaa3… jyu4gwo2 jyu4 mou4 ji3ngoi6.

the-Great-Wall SFP if if NEG accident

‘This one is obvious. It should be the Great Wall of China… if nothing goes wrong.’

As observed in (4), the sentence final particle gaa3 and wo3 are used as devices to upgrade the speaker’s epistemic stance and to make the reason as to why it is not China stronger. On the other hand, gwaa3 in (5) is rather considered as an epistemic downgrader, indicating the uncertainty of the speaker. When being asked where the place is, the interviewee prefaces his answer with Ni1 go3 hou2 ming4bin2 laa1 ‘this one is obvious’, reinforced by the affirmative marker laa1. However, as soon as he has said that, he probably realises that the claim just made is too strong and thus he shifts towards the other end of the scale of epistemic strength. The particle gwaa3, or even the afterthought jyu4gwo2 jyu4 mou4 ji3ngoi6 ‘if nothing goes wrong’, can be seen as strategies to mitigate his claim.

As the two examples demonstrate, the speaker’s strategic choice of a particular particle, whether it is a gaa3, a wo3, or a gwaa3, has important consequences for the overall epistemic strength that the speaker attempts to achieve.
4.1.4 Discourse-pragmatic strategies

Besides the use of epistemic/evidential markers and sentence final particles, speakers of Cantonese can also employ a wide range of discourse-pragmatic strategies to modulate, particularly to downgrade, the degree of epistemic strength of their utterance. These strategies include the use of hedging devices and redressive language, keeping silent, as well as providing a reason/justification.

(6)

01 IR: Nei5 tai2-ba/6 jing6-m4-jing6-don2 bin1 jat1 go3 dei6fong1

2SG look-ASP recognise-NEG-recognise-RVC which one CL place

ni1don6 hai6?

Here COP

‘Take a look and see if you can recognise which place it is.’

02 IE: (6.0) Hai6-mai6 Hon4Gwok3 ding6 Jat6Bun2 go2 di1 aa3?

COP-NEG Korea CONN Japan that CL Q

‘Is that Korea or Japan or the like?’

(7)

01 IR: Gam2 nei5 gok3dak1 hai6 ci5 Zung1Gwok3?

Then 2SG feel COP seem China

‘Then do you think it seems like China?’

02 IE: Hai6, hai6.

Yes yes

‘Yes, Yes.’

03 IR: Dim2gaai2 ci5 Zung1Gwok3 aa3?

~ 118 ~
Why seem China Q

‘Why does it look like China?’

04 IE: e4…e4zik1-hai6 e4…jan1wai4 keoi5 ni1 di1 gin3zuk1 e4…jan5 di1

uhuh that-is uh because 3SG this CL building uh EXIST CL
ci5 go2 di1 z6min2 waak6ze2 ting4jyun2 go2 di1 gam2gok3.
resemble that CL temple or courtyard that CL feeling

‘Uh…uh…that is…uh…because I feel that such building is pretty like
those temples or courtyards.’

The markedly long silence in (6) indicates that the speaker is probably very unsure
about what the place is, and the use of the indefinite pronoun go2 di1 ‘those’ also
echoes the speaker’s uncertainty. Turning to (7), we observe that a number of
discourse-pragmatic strategies have been used to downgrade the degree of
epistemic commitment of the speaker, as shown in line 4. The repeated use of the
hedging e4 ‘uh’, together with the gap filler zik1hai6 ‘that is’ and the deliberately
redressive language used as indicated by the numerous indefinite go2di1 ‘those’,
which might also be considered a kind of gap fillers here, the speaker has shown
his lack of confidence in his answer to the interviewer’s question, even though he
is attempting to give an explanation.

4.2 Modulating strength in conversations

The examples above have demonstrated how each category of devices work in the
discourse-pragmatic context of an interview. In this section, we will look at some
longer stretches of discourse excerpted from our data to see how various
strategies can collaborate with one another in an interactional context.
01 IR: Ni1 jat1 go3 nei5 gu2-m4-gu2-dou2 hai6 bin1dou6?
   This one CL 2SG guess-NEG-guess-RVC COP where
   ‘For this one, can you guess where it is?’

02 IE: Ni1 go3 hai6 Zung1Gwok3 lo1.
   This CL COP China SFP
   ‘For this one, it is (obviously) China.’

03 IR: Zung1Gwok3? Dim2gaai2 gok3dak1 hai6 Zung1Gwok3 ge2?
   China why feel COP China Q
   ‘China? Why do you think it is China?’

04 IE: Jing1goi1 hai6 Zung1Gwok3 ge2, m4-zi1 wui6-m4-wui6 hai6
   MOD COP China SFP NEG-know MOD-NEG-MOD COP
   Zung1Gwok3 ne1? Jan1wai4 keoi5 jan5 di1 Zung1Man4 zi6 ge3
   China Q because 3SG EXIST CL Chinese character GEN
diu1bak1 hai2 dou6.
   sculpture at here
   ‘It should be China. I’m not sure if it is China. Because there are sculptures
   on which some Chinese characters are written.’

05 IR: Bat1gwo3 ngo5 soeng2 man6 do1 nei5 sin2-si2… dim2gaai2 nei5 gok3dak1
   CONN 1SG want ask more 2SG little-little why 2SG feel
   jan5 Zung1Man4 zi6 zan6 hai6 Zung1Gwok3 ne1?
   EXIST Chinese character then COP China Q

06 IE: 4(laughing)…jan1wai4…jan1wai4…tung1soeng4 dou1 hai6… zik1-hai6
   uh (laughing) because because usually all COP that-is
   jyu4gwo2 gu2zik1 go2 di1 je5 soeng6-min6 jan5 bak1-zyu6
if monument that CL thing upside EXIST sculpture-ASP
Zung1man4 zil6, ngo5 nam2 jing1goi1 dou1 hai6 wui6 hai6
Chinese character 1SG think MOD also COP MOD COP
Zung1Gwok3 jap6-min6 go2 di1 gu2zjk1 gam2 joeng6 laa1.
China inside that CL monument such look SFP
‘Uh (laughing)... because...because...usually... I mean, if Chinese characters appear on the sculpture, I think it is possible that it would be the monument in China.’

07 IR: Dim2gaai2 m4-hai6 Jat6Bun2 m4-hai6 Hon4Gwok3 aa3?
Why NEG-COP Japan NEG-COP Korea Q
‘Why isn’t it Japan or Korea?’

08 IE: Hai6 wo3, gam2 nei5 jau6 man6-dak1-ngaam1 wo3! Ngo5 dou1 kei4sat6
Right SFP such 2SG also ask-PRT-right SFP 1SG also actually
dou1 bo2nang4 hai6 sin1jap6wai4zyu2 wo3... (laughing)... kei4sat6 ngo5 dou1
also MOD COP subjective SFP (laughing) actually 1SG also
jau5 nam2-gwo3 hai6-mai6 Jat6Bun2. Jan1wai6 Jat6Bun2 dou1 hou2 do1
EXIST think-ASP COP-NEG Japan because Japan also very many
go2 di1 e4 zil6min2 aa3 go2 di1 hai6 gam2 jeong2 gaa1 ma3.
that CL uh temple SFP that CL COP this look SFP SFP
‘Right. You have asked a good question! I may be actually subjective...(laughing)... Actually I have thought if it is Japan, because there are also many temples in Japan look like this one.’
In (8), we can observe that the interviewee has made use of a wide range of strategies to gradually downgrade her epistemic commitment of the claim she made. In line 2, she makes a strong claim that \textit{it IS China}, \textit{hai6 Zung1Gwok3 lo1 BE-China-SFP}, reinforced by the copula \textit{be} and the affirmative particle \textit{lo1}. However, she gradually mitigates this claim in her subsequent turns (i.e. in lines 4, 6, and 8). In line 4, the interviewee makes use the epistemic modal \textit{jing1goi1} ‘should’ and the mitigative sentence final particle \textit{ge2} in conjunction with her claim \textit{hai6 Zung1Gwok3 BE-China}. She further weakens her claim by actually questioning her own claim using \textit{m4-zi1 wui6-m4-wui6} ‘not sure if …’, even though she is trying to provide a reason for why still it is highly possible that it is China (i.e. the fact that there are Chinese words). After being further challenged by the interviewer in line 5 as to why having Chinese words means China, the interviewee continues to downgrade her claim. The latter part of line 6 has already been discussed in example (1), when we talked about epistemic modals and adverbials. In addition to the psych verb \textit{nam2} ‘think’, the modal auxiliary \textit{jing1goi1} ‘should’ and \textit{wui5} ‘would’, which we have discussed earlier, the hedger \textit{e4} ‘uh’ and the use of conjunctions such as \textit{jan1wai6} ‘because’ and \textit{zik1hai6} ‘that is’ (both carry a discourse function of a gap filler), together with the embarrassed laugh, they work together to claim a lower degree of epistemic commitment to the claim originally reported. When the interviewer further challenges the interviewee in line 7, by actually proposing \textit{Jat6Bun2} ‘Japan’ and \textit{Hon4Gwok3} ‘Korea’ as two possibilities, the interviewee steps back in line 8, making a concession that the interviewer might in fact be right, as Japan and Korea also have many temples similar to the one depicted in the picture. By using \textit{kei4sat6} ‘in fact/actually’ after the embarrassed laugh, the speaker proposes a revised conclusion and attempts to align with the interviewer’s proposal that it might in fact be Japan.

(9)

01 IE: Ni1 go3 cit3gai3 a3 Dak1Gwok3 gn2bou2.

This CL design seem German castle

‘This one looks like a German castle.’
02 IR: Ni1 go3 ci3 hai6 di1 Tin1Zyu2Gaau3-sik1 go2 di1 gin3zuk1.

This CL seem COP CL Catholic-style that CL architecture
‘This one looks like those catholic architectures.’

03 IE: Ji2, gam2 nei5 jau6 gong2-gong2-haa5 jau6 hon2ci3 hai6=daan6hai6 keoi5
INJ then 2SG again talk-talk-ASP again seem COP but 3SG
don1 mou5 sap6zi6gaa2.
also NEG crucifix
‘Well after hearing what you said it seems like – but it doesn’t have a crucifix.’

04 IR: Gam2 keoi5 m4-jat1ding6 jin3 hai2 ni1 go3 wai2 tai2 don2 sap6zi6gaa2
Then 3SG NEG-MOD MOD at this CL place see-RVC crucifix
ge3 ze1=daan6hai6 ngo5 gok3dak1 ni1 go3 wai2 ci3 lo1, nei5
SFP SFP but 1SG feel this CL place seem SFP 2SG
gok3dak1 ne1?
feel Q
‘Well but the crucifix doesn’t need to be seen from here, but I think it looks like (a catholic architecture) from this angle, what do you think?’

05 IE: m6 … (4.0) ngo5 gok3dak1 (1.0) m4-hai6 taai3=zik1-hai6 ci5
Hmm 1SG feel NEG-COP see that-is resemble
ge3 daan6hai6 jyu4gwo2 nei5 … ne1… tai2-jaan1 seng4 fuk1 tou4
SFP but if 2SG PRT see-PRT whole CL picture
ngo5 wui3 gok3dak1 keoi5 go3 gin3zuk1 ci3 Dak1Gwok3 gn2bon2
1SG MOD feel 3SG CL architecture seem German castle
do1 di1.

~ 123 ~
more CL.

‘Hmm… I think… not so … it does look like (a catholic architecture) but if you go back to the picture I would still think that this architecture looks like a German castle more’

Example (9) above is another instance in which the interviewee progressively weakens her degree of epistemic commitment to the claim made. She first makes a suggestion in line 1 that the picture shows a German castle. The interviewer makes a counter suggestion in line 2 that the building looks like a catholic architecture. In here, the interviewee starts to step down a bit in line 3 by attempting to ratify the interviewer’s view, though she possesses a challenge that the building does not seem to have a cross. In line 4, the interviewer steps up again and reinforces her claim by using the particle ze1 to increase her aggressiveness to ask the question. In line 5, we observe that the interviewee has become more reluctant and has adopted a range of strategies to modulate her claim. The two silences indicate that the speaker probably needs to buy time to think of ways to support her claim. In addition, the m6 (as both hedger and gap filler), the epistemic phrase ngo5 gok3dak1 ‘I think’, together with the implicit evidential marker ci3 ‘seems’ and various conjunctions such as zik1hai6 ‘that is’, daan6hai6 ‘but’, and jyu4gwo2 ‘if’ which carry the discourse-pragmatic function of gap fillers and hesitators to mitigate the epistemic stance of the interviewee, all work together in a collaborative manner to downgrade the interviewee’s epistemic commitment to the claim.

4.3 Epistemic modulations and politeness considerations

From the data above it is observed that epistemic downgrading strategies are more frequently used than upgrading strategies. An equally worth-noting observation is that when epistemic upgrading is involved, it is mostly done by the use of sentence final particles, as illustrated in example (4) in section 4.1.3. We believe that there are politeness implications behind these observations. It is always considered impolite when one shows disagreement because this could
damage the other person’s positive face. The tacit rule that people in a Chinese community generally agree to conform is that we should observe each other’s face in the course of a conversation, and try not to damage their face needs. When a speaker needs to upgrade his claim, he would choose to do it in a more implicit and subtle way. This is why sentence final particles are always employed as the grammatical strategy for epistemic upgrading in conversations.

In the Chinese culture, people are expected to use language in a certain way so that their behaviour is considered polite. Previous studies on politeness phenomena in Chinese seem to suggest that politeness strategies are face-driven in the Chinese culture (Gu 1990; Zhan 1992; Pan 2002; Chen 2003; among others). Chinese people should respect one another and attend to the face needs of other people in order to be polite. “Face” is very important to Chinese people, as the saying min6zi2 daai6-gwo3 tin1 (lit. face bigger sky ‘one's face is bigger than the sky’) suggests. It is considered very impolite if one does not take into consideration other people’s face needs. Thus, in response to the challenges or disaffiliative responses from their addressees, speakers would modulate, in particular downgrade, the epistemic strength of their claims so that they observe the face of their addressees. The various downgrading strategies that we have discussed in this paper are doing a lot of face work in the course of a conversational interaction. Speakers can many a time downgrade their claims so as to elicit alignment and attain common ground with their addressees. As observed in line 8 of example (8) and line 3 of example (9), the speaker is backing down and making a concession, by even offering a revised conclusion in the former example. In line 4 of example (8), the act that the interviewee actually questions her original claim is considered to be doing some intersubjective work. In fact, attempting to involve another party in the conversation is also considered an act of face-saving.

Unlike the Chinese culture, which is face-driven, the Anglo culture is knowledge-driven. The Anglo culture has a value and an emphasis on accuracy and thus it is

---

7Positive politeness refers to the redress to the addressee's positive face, which is one's want that his/her wants be desirable to at least some others. An act of disagreement might damage the positive face of the addressee as it does not fulfil the addressee's want that his/her opinion be desirable. See Brown & Levinson (1978) for the details on the politeness phenomenon.
closely linked with the use of “understatement” (Wierzbicka 2006). From the Anglo cultural point of view, one cannot say all if he actually means “some” or “many”, or say none if he actually mean “a few”. One should not exaggerate, but rather should say exactly what one wants to say (i.e. no more and no less). Wierzbicka sees the rise of epistemic phrases as one of the most remarkable phenomena in the history of the English language that reflects the practice of “understatement” in English. Epistemic phrases such as I think, I believe, I guess, and I assume are phrases clarifying the speaker’s stance in relation to what is being said (ibid: 204). For example, declarative sentences like “Bill is in Sydney now” imply an unqualified claim to knowledge: “I know.” By contrast, sentences with epistemic phrase like I think, I guess, or I assume (e.g. “I think Bill is in Sydney now”, “I guess Bill is in Sydney now”, “I suppose Bill is in Sydney now”) do not imply any such claim.

As Wierzbicka remarks, one would be struck by the wealth of epistemic phrases (or epistemic “downtoners”) like I think, I guess, and I assume when one looks at the daily interactional use of English. The major function of these phrases is to acknowledge the limitation of one’s knowledge and to indicate the tentativeness of most of the things that s/he says, which echoes the “understatement” practice which is agreed among speakers of English. They do not say anything which lacks evidence or a reason. The Anglo culture would expect that speakers can make use of language to acknowledge the limitations of their knowledge about the world.

5. Summary and conclusion
This study expands the scope of studies of evidentiality and epistemic modality, as aspects of the stance-taking research, by exploring the strategies that speakers use to calibrate the epistemic strength of their utterances. Participants in a conversation mutually observe each other’s shift of stance and change in epistemic commitment to the given information. As the findings of the current investigation clearly show, epistemic stance is not something that is absolute or fixed, but rather is something being continuously shaped and negotiated among the participants of a talk. In this paper, we have shown how speakers of Cantonese can deploy various strategies to modulate, in particular to downgrade,
their epistemic strength. These strategies include the use of epistemic modals and adverbials, evidential markers, sentence final particles, as well as a wide range of discourse-pragmatic strategies, such as the use of silences and hedging devices.

Stance-taking as a pragmatic phenomenon is complex and multi-faceted. As an area of study, it challenges us with its many layers of complexity, which prior studies have continued to untangle from different perspectives. More studies from an interactional or discourse perspective like the present one will surely advance our understanding of the intricacy of our everyday conversational communication of expressing and negotiating stance. Through our analysis of stance-taking strategies in Cantonese, we also hope to shed light on other cross-linguistic studies on evidentiality, epistemicity, and stance-taking.

Acknowledgements
The author wishes to gratefully acknowledge the generous support of the Research Grants Council of Hong Kong for the project “Epistemic Modulation and Speaker Attitude in Cantonese: A Discourse-pragmatic Perspective” (UGC/FDS16/H07/14). The author also wishes to thank the audience at ALS 2014 for their valuable feedback, the anonymous reviewer for the very helpful comments and suggestions, and Nikko Yin-yee Lai for assisting with data collation and transcription. Any shortcomings in this paper are the author’s own responsibility.

References


~ 128 ~


Leung, C. S. (2005). *A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.


Predatory chatlog messages as forensic evidence in court:

A comparison of two different procedures for estimating the weight of evidence

Shunichi Ishihara
The Australian National University
shunichi.ishihara@anu.edu.au

Abstract. The performances of two different procedures for calculating the likelihood ratio (LR) for forensic text comparison (FTC) are empirically compared. One is the multivariate kernel density (MVKD) procedure with so-called lexical features. The MVKD procedure has been successfully applied to various types of evidence, including texts. The other is the procedure based on character $N$-grams. $N$-gram is a widely-used, robust probabilistic language model. The effectiveness of character $N$-grams has been reported in authorship analysis, however, to the best of my knowledge, it has not yet been applied to LR-based FTC. In this study, the log-likelihood-ratio-cost ($C_{llr}$), which is an appropriate assessment metric for LR-based systems, is used to assess the performance of the two procedures. It will be reported that the MVKD procedure outperforms the character $N$-gram procedure. Through the comparison of the two procedures, this study also demonstrates how the weight of evidence (= LR) can be estimated from text messages.

Keywords: likelihood ratio, chatlog messages, multivariate kernel density, character-based $N$-grams, log-likelihood ratio cost
1. Introduction

Forensic text comparison (FTC), which has been traditionally called forensic authorship analysis, is a new area of study, in which by analysing the offender's and suspect's text samples, a forensic scientist assists the trier-of-fact in deciding whether they are from the same author (the suspect is guilty) or different authors (the suspect is exonerated). The application of authorship analysis to the field of forensics is relatively new, yet authorship analysis itself is a long established area of study which emerged first as stylometric studies\(^1\) (Holmes, 1992; Mendenhall, 1887, p. 64; Mosteller & Wallace, 1984; Thisted & Efron, 1987).

In February, 2015, an article entitled “Should texts, e-mail, tweets and Facebook posts be the new fingerprints in court?” was published on the Washington post\(^2\). It is evident from this article that subjective approaches with qualitative data analysis have widely been employed in authorship cases, despite that the necessity of objective and quantitative approaches has been fervently argued in forensic comparative sciences in general (Saks & Koehler, 2005).

In the current study, I apply the same framework used in forensic DNA analysis, namely the likelihood ratio framework of evidential strength (henceforth, the LR framework) to predatory chatlog messages so as to estimate their evidential strength. Researchers and forensic scientists engaged in forensic authorship analysis are well aware of LR and its importance in forensic comparative science. For example, the word ‘LR’ appears many times in papers, included in the 2\(^{nd}\) issue of volume 21 of *Journal of Law and Policy*, which was published in 2013 as the proceedings of the papers presented at a forensic authorship attribution workshop\(^3\) held in October 2012. However, contrary to its importance, LR-based studies on forensic authorship analysis are conspicuous in their rarity (cf. Ishihara, 2013, 2014a, 2014b). It is not surprising that the LR framework was not mentioned in the article of the Washington post obviously because it is not a standard framework in authorship cases.

---

\(^1\) Stylometry is the science of measuring literary style.

\(^2\) https://www.washingtonpost.com/lifestyle/magazine/should-texts-e-mail-tweets-and-facebook-posts-the-be-new-fingerprints-in-court/2015/02/19/a5cc2bf6-6f32-11e4-8808-afaf1c3a33cf_story.html

\(^3\) http://www.brooklaw.edu/newsandevents/events/2012/10-11-2012a.aspx
In this study, I use the so-called ‘character-based’ N-grams to statistically model each author’s attribution, and compare the result with that of another procedure – the multivariate kernel density (MVKD) procedure (Ishihara, 2014b)— to see which procedure performs better. The MVKD procedure has been successfully applied to various types of forensic evidence, including handwriting (Bozza, Taroni, Marquis, & Schmittbuhl, 2008; Hepler, Saunders, Davis, & Buscaglia, 2012), fingerprint (Neumann et al., 2007), voice (Gonzalez-Rodriguez, Rose, Ramos, Toledano, & Ortega-Garcia, 2007; Morrison, 2009), texts (Ishihara, 2014b) and more. The effectiveness of character N-grams has been reported in authorship analysis (R. Layton, Watters, & Dazeley, 2012; Stamatatos, 2013), however, it has not yet been applied to LR-based FTC. Thus, we don’t know which procedure works better in FTC. In this study, the performances of the above-mentioned two procedures are objectively compared using the same text data. The performance of the MVKD procedure is reported in Ishihara (2014b), and it is compared with the performance of the character N-grams in this study.

Since there are various procedures to estimate forensic LRs (Aitken & Lucy, 2004; Doddington, 2001; Lindley, 1977; Nair, Alzqhoul, & Guillemin, 2014; Reynolds, Quatieri, & Dunn, 2000), it is important to empirically test the performances of them one by one. Through the comparison of the two procedures, this study also demonstrates how the weight of evidence (= LR) can be estimated from electronically generated text messages.

2. Bayes’ Theorem and Likelihood Ratio

An FTC typically involves the comparison of text samples of known origin (e.g. the suspect’s text sample) with other text samples of disputed origin (e.g. the text samples of the offender). The FTC expert will employ certain statistical techniques on data extracted from the text sample evidence with the ultimate aim of assisting the trier-of-fact in their final decision. When the trier-of-fact is faced with the task of making the final decision of the suspect being guilty or not guilty, the trier-of-fact needs to consider two conditional probabilities; the probability of the suspect and the offender being the same individual, and the probability of the suspect and the offender being different individuals, and these two probabilities are conditional on the evidence. The ratio of these two probabilities is known as the posterior odds in Bayes’ Theorem, and can be mathematically expressed as follows:
1) \[ \frac{p(H_p|E)}{p(H_d|E)} \]

In 1), \(p(H_p|E)\) represents the probability of the prosecution hypothesis \((H_p)\), given \((E)\) the evidence \((E)\) (the probability that the suspect and the offender are the same person given the evidence), and \(p(H_d|E)\) represents the probability of the defence hypothesis \((H_d)\) given the same evidence (the probability that the suspect and the offender are different people given the same evidence). If \(p(H_p|E)\) is 90% and \(p(H_d|E)\) is 10%, for example, the ratio between them is 9. It means that given the evidence, it is 9 times more likely that the suspect and the offender are the same person than they are different people. However, if the FTC specialist were to provide the trier-of-fact with this strength-of-hypotheses statement (as opposed to a strength-of-evidence statement), they would in effect be making a statement about the suspect’s guilt or innocence, which is usurping the privilege belonging only to the trier-of-fact (Aitken, 1995; Evett, 1998). By Bayes’ Theorem given in 2), the task of the trier-of-fact and that of the forensic expert can be clearly separated.

2) \[ \frac{p(H_p|E)}{p(H_d|E)} = \frac{p(H_p)}{p(H_d)} \cdot \frac{p(E|H_p)}{p(E|H_d)} \]

As seen in 2), the posterior odds is the product of the prior odds and the likelihood ratio. The posterior odds and prior odds belong to the province of the trier-of-fact and the LR to the province of forensic experts. That is, Bayes’ Theorem and, by extension, the LR framework enable the forensic expert to make an objective strength-of-evidence statement, and in doing so, they can avoid usurping the role of the trier-of-fact. Bayes’ Theorem also outlines the appropriate way for updating belief (prior odds) in the two hypotheses as new evidence is introduced. Furthermore, the forensic expert cannot logically express conclusions (suspect is guilty or not) in the form of posterior odds as they are not privy to the prior odds and they cannot make access to the analyses of other evidence in the case.

Thus, as explained above, the task of the forensic scientist is to estimate the strength of evidence, which is technically referred to as Likelihood Ratio (LR).
Put simply, the LR is the probability of the evidence occurring if an assertion were true, divided by the probability that the evidence would occur if the assertion were not true (Robertson & Vignaux, 1995, p. 17). In typical FTC analyses, two sets of text samples are analysed. FTC analysis is carried out on the assumption that one set of text samples is from a known author (e.g. suspect), and the other is from an unknown author (e.g. offender). Thus, the LR given in 2) can be interpreted as the probability ($p$) of observing this ‘evidence’ ($E$ = the particular differences between the two sets of text samples) given that the prosecution hypothesis ($H_p$) is true (the suspect and the offender are the same person), relative to the probability ($p$) of observing the same evidence ($E$) if the defence hypothesis ($H_d$) is true (the suspect and the offender are different people). For example, a calculated LR of 20 means that the evidence is 20 times more likely to arise if the text samples are from the same author, than if the text samples are from different authors. Note that this is not the same thing as saying that it is 20 times more likely that the two sets of text samples are of the same author than different authors, which would be the interpretation of a posterior odds of 20.

The LR given in 2) is always expressed as a positive value. The greater the LR is over 1, the stronger the support for the prosecution hypothesis (the suspect and offender are the same author), and the smaller the LR is below 1, the stronger support for the defence hypothesis (the suspect and offender are different authors). That is, the amount of the deviation from LR = 1 quantifies the strength of the evidence. It is important to emphasise that the LR is not a binary expression of truth; it does not give an answer to the question ‘were these two text samples written by the same person or different people?’, but rather it assesses how strongly the evidence relatively supports the hypotheses.

The important point is that the LR is concerned with the probability of the evidence, given the hypotheses, which is the province of forensic scientists, while the trier-of-fact is ultimately concerned with the probability of the hypothesis, given the evidence.

LR values are often expressed in logarithmic form (e.g. log\(_{10}\)). This is because the log LR is intuitively easier to understand; $LR = 10$ (log\(_{10}\)LR = 1) and $LR = 0.1$ (log\(_{10}\)LR = -1) hold the same strength although they support opposite hypotheses.
3. Experiment design

As stated in §1, the aim of this research is to compare the performances of two different procedures for estimating LRs, namely the MVKD based and character N-gram based procedures. These two procedures are compared under exactly the same conditions (e.g. the use of database and assessment metrics). Although the detailed descriptions of the experiments and the MVKD procedure are provided in Ishihara (2014b), they are concisely given in this section for the convenience of the reader. However, the reader is still encouraged to go through Ishihara (2014b). Thus, a large part of this section is spared for the explanation of the character N-grams (§3.2.2).

3.1 Databases

For the experiments, chatlog messages written by 115 authors were randomly selected from an archive of chatlog messages⁴, which is a collection of real pieces of chatlog evidence used to prosecute paedophiles. These 115 authors were all convicted between years 2007 and 2012. They are from 28 different states of America⁵. Their average age is 37.4 years old (sd = 11.4; min = 19 and max = 66).

In order to assess the performance of an FTC system, two types of comparisons, namely same-author (SA) and different-author (DA) comparisons, are necessary. In SA comparisons, two groups of messages produced by the same author are compared and evaluated with the derived LR. Given the same origin, it is expected that the derived LR is higher than 1. In DA comparisons, mutatis mutandis, they are expected to receive an LR lower than 1. The messages belonging to each of the 115 authors were separated into two non-contemporaneous⁶ groups so that SA and DA comparisons can be carried out by simulating an offender-suspect situation.

⁴ http://pjfi.org/

⁵ The break-downs of these 28 states are: Michigan = 28; California = 17; Kentucky = 8; Georgia and New Jersey = 7; Florida = 6; Utah and Indiana = 4; Alabama, Connecticut, Oklahoma, Oregon and Texas = 3; Idaho, Nevada, New York, Ohio, Pennsylvania, Tennessee and Wisconsin = 2; Arizona, Arkansas, Louisiana, Maine, Mississippi, North Carolina, North Dakota and Washington = 1.

⁶ The reviewer pointed out that the importance of non-contemporaneity is not clear for text to the same extent that it is to speech. To the best of my knowledge, there has not been any studies which investigated to what extent the non-contemporaneity of texts furnishes within-author variability. However, considering forensically realistic conditions of FTC and erring on the side of caution, two non-contemporaneous groups of messages were selected from each author for the current study.
The 115 authors were further divided into mutually-exclusive test (39 authors), background\(^7\) (38) and development (38) databases. The test database is used to assess the performance of the FTC system, the background database is used as the reference database (in terms of typicality) for calculating LRs, and the development database is to calculate weights for calibrating the derived LRs (Refer to §3.3 for calibration). From the test database of 39 authors, 39 independent SA and 1482 DA comparisons are possible\(^8\).

I use 500 words in order to model the authorial attributes of each message group. Note that the previous paragraph contains 145 words; 500 words is c.a. 3.5 times as many words as the paragraph. Recent advances in text mining have drastically enhanced the capability of authorship analysis even with limited amount of data. For example, Robert Layton, Watters, and Dazeley (2010, p. 1) demonstrated that authorship identification can perform “at rates significantly better than chance for documents of 140 characters or less”. However, to the best of my knowledge, there have not been any comprehensive studies describing how the amount of data influences the performance of LR-based FTC systems, in particular, in terms of the magnitude of the derived LRs.

3.2 Estimation of likelihood ratio: The multivariate kernel density procedure and the character \(N\)-gram procedure

This subsection explains how messages are modelled and how the LRs are estimated in the two procedures in question. A concise explanation of the MVKD procedure, including the features used, is given in §3.2.1 for the convenience of the reader. A detailed description of the character \(N\)-gram procedure is given in §3.2.2.

\(^7\) A background sample is used to model the distribution of authorship attribution properties in the relevant population, which is determined by the defence hypothesis of each case. The relevant population is the population to which the true perpetrator of the crime could have conceivably belonged (Ochoa & Morrison, 2011, p. 64). Refer to Aitken and Taroni (2004) and Robertson and Vignaux (1995) for detailed discussion of the relevant population mainly in the context of DNA. Also refer to Ochoa and Morrison (2011) and Gold and Hughes (2014) for their arguments of practical procedures for collecting samples from the relevant population for forensic voice comparison.

\(^8\) As for DA comparisons, two independent different DA comparisons are possible (e.g. (A)uthor1(G)roup1 vs. A2G1 and A1G2 vs. A2G2) for each pair of different authors (e.g. A1 vs A2). Since the number of non-overlapping different author pairs is 741 (= 39C2) and two independent different DA comparisons are possible for each different pair, as a whole, 1482 DA comparisons (= 741*2) are possible (39C2*2).
3.2.1 Multivariate kernel density procedure and lexical features

Following the results of previous authorship studies (De Vel, Anderson, Corney, & Mohay, 2001; Iqbal, Binsalleh, Fung, & Debbabi, 2010; Ishihara, 2012; Zheng, Li, Chen, & Huang, 2006), the features listed in Table 2 were used in Ishihara (2014b) in order to model each message group. These features are, in a broad sense, lexical features. All of these features are the continuous features of numerical values. These features are vocabulary richness (1~3), token-based (4~5) and character-based (6~12) features. That is, the authorship attribute of a message can be modelled or represented as a vector of as many as 12 feature values \( m = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{12}\} \). One feature vector was obtained from each message group – each author has two message groups. The features listed in Table 2 were used in the current study in order to model each message group.

<table>
<thead>
<tr>
<th>Type</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>vocabulary</td>
<td>1. Yule’s I (the inverse of Yule’s K)</td>
</tr>
<tr>
<td>richness</td>
<td>2. Type-token ratio (TTR)</td>
</tr>
<tr>
<td></td>
<td>3. Honoré’s R</td>
</tr>
<tr>
<td>lexical:</td>
<td>4. Average token number per message line</td>
</tr>
<tr>
<td>token-based</td>
<td>5. SD of the token number per message line</td>
</tr>
<tr>
<td>lexical:</td>
<td>6. Average character number per message line</td>
</tr>
<tr>
<td>character based</td>
<td>7. SD of the character number per message line</td>
</tr>
<tr>
<td></td>
<td>8. Upper case character ratio</td>
</tr>
<tr>
<td></td>
<td>9. Digit character ratio</td>
</tr>
<tr>
<td></td>
<td>10. Average character number per token</td>
</tr>
<tr>
<td></td>
<td>11. Punctuation character ratio ( (, . ? ! ; : ’ ”) )</td>
</tr>
<tr>
<td></td>
<td>12. Special character ratio ( (&lt;&gt; %</td>
</tr>
</tbody>
</table>

Table 2. List of features (from Ishihara (2014b, p. 48)).

Experiments were repeatedly carried out by changing the length of a feature vector (or using different feature sets) so as to investigate which feature combination results in the best performance. Only some feature combinations were tested in Ishihara (2014b), but they were done in a systematic manner. First of all, all possible combinations of two features \( \{f_{so}, f_t\} \) were tested, and the five best performing bi-features were selected. Using these five best performing bi-features as bases, the performance of the tri-features \( \{f_{so}, f_{lo}, f_{lt}\} \) was tested by adding one of the remaining features one by one to these bases. This process was repeated for feature vectors of higher dimensions.
With the features given in Table 2, LRs were estimated by means of the MVKD procedure (Aitken & Lucy, 2004). Although the reader needs to refer to Aitken and Lucy (2004) for a full mathematical exposition of the MVLR formula, it accepts multiple continuous features, such as given in Table 2, as input parameters and takes their correlations into account to estimate LRs. The MVKD procedure was successfully adopted by various forensic comparative sciences for estimating the weight of evidence, including handwriting (Bozza et al., 2008; Hepler et al., 2012), fingerprint (Neumann et al., 2007) and voice (Gonzalez-Rodriguez et al., 2007; Morrison, 2009). The MVKD procedure assumes normality for within-group (e.g. within-author) variance while it uses a kernel-density model for between-group (e.g. between-author) variance. The MVKD procedure bases the estimate of the within-group and between-group distributions on a background database.

Appendix 1 contains the normal QQ plots for the twelve features given in Table 2. As can be seen in these normal QQ plots, none of the twelve features exhibits normality in their distributions. Appendix 2 is a matrix of correlation coefficients between the twelve features. It is apparent from these that many of the features show certain degrees of correlations with another feature. These are the main motivations for the use of the MVKD procedure for the lexical features given in Table 2.

Theoretically speaking, the LRs estimated by the MVKD formula should be well-calibrated. However, when the modelling assumptions of the formula are violated, the resultant poorly calibrated LRs estimated by the MVKD formula, which are customarily referred to as scores, need to be calibrated (refer to §3.3 for a detailed explanation of calibration) before they are interpretable.

3.2.2 Character N-grams

In authorship analysis, a message can be viewed as consisting of a series of items, such as characters, words, sentences, etc. and the characteristics of the message can be represented or modelled by a series of these items. An N-gram model was developed as a type of probabilistic language model for predicting the next item(s) in a sequence (Jurafsky & Martin, 2009). It is considered that the character N-grams can capture not only information at a character level (e.g. spelling characteristics) but also even the information of higher linguistic levels (e.g. formality from the use of punctuation marks).
If the author is an American person who is relatively casual in writing messages, the sequence of characters, such as ‘zed’ (not ‘sed’ as in ‘realised’), ‘(-:’ or ‘lol’, may more frequently appear in his messages than for messages written by an Australian whose writing style is relatively formal. One way of quantifying these types of characteristics (e.g. ‘zed’, ‘(-:’ or ‘lol’) in a probabilistic manner is to consider the probability of the sequence of these characters. That is, in the case of ‘sed’ as an example, it is the probability of the character ‘d’ following the sequence of ‘s’ and ‘e’. Formally, this probability can be written as $p(d|s, e)$ – given the sequence of ‘s’ and ‘e’, the probability of the character ‘d’. The same sequence (‘sed’) can be dealt with in terms of a sequence of two characters. In this case, it is the probability of the character ‘e’, given ‘s’ ($= p(e|s)$) and the probability of the character ‘d’ given ‘e’ ($= p(d|e)$), and the overall probability of the same sequence (‘sed’) is the product of them ($= p(e|s) * p(d|e)$), assuming that these are independent. The $N$ of character $N$-grams means the number of characters constituting the sequence in question; $N = 1$ is called unigram (e.g. $p(s)$), $N = 2$ is bigrams (e.g. $p(s|e)$), $N = 3$ is trigram (e.g. $p(d|s, e)$), and so on. The actual probability of, for example, a sequence of two characters $C_{n-1}, C_n$ (e.g. ‘se’) is estimated by dividing the count of the sequence $C_{n-1}, C_n$ (e.g. ‘se’) by the count of $C_n$ (e.g. ‘e’) in the data (count($C_{n-1}, C_n$)/count($C_n$)).

For the sentence ‘forensic analysis is important’, for example, the character ‘i’ appears 4 times and the character sequence of ‘si’ 2 times. Thus, $p(i|s)$ is 0.50 ($= 2/4$). Furthermore, the 3 occurrences of the character ‘n’ and the 2 occurrences of the character sequence of ‘an’ will result in $p(n|a)$ as 0.66 ($= 2/3$). In this manner, we can model a message with a series of character $N$-grams and their associated probabilities. If a given message (e.g. an incriminating message written by an offender) is similar to another given message (e.g. a message written by a suspect), the degree of the similarity between them should be reflected to the probability of the incriminating message against the character $N$-gram model of the suspect message.

In the current study, the formula proposed by Doddington (2001) is used for estimating LRs because this formula is specifically designed to calculate LRs based on $N$-grams. Doddington’s (2001) formula is given in 3). The output of the formula is a score, not an LR, and the score later needs to be converted to an LR by means of calibration (refer to §3.3 for calibration).
Consider the situation whereby two messages (e.g. \(i\) and \(j\)) need to be compared and analysed for the LR. In an \(N\)-gram based procedure, the first process is that one of the messages (e.g. \(i\)) needs to be modelled, and have the calculation of the \(N\)-gram probabilities for the message \((i)\). In 3), the model of the message \((i)\), which consists of a list of the \(N\)-gram probabilities, is denoted as \(A_{\text{author}}^i\). Then the degree of similarity between the two messages \((i\) and \(j)\) can be quantified by probabilistically estimating how well the other message (e.g. \(j\)) fits the model of the first message \((A_{\text{author}}^i)\). This process of assessing the similarity of two messages \((i\) and \(j)\) is denoted as \(A_{\text{background}}^j\) in 3). In order to estimate scores or LRs, besides the similarity of the two messages, the typicality needs to be assessed, which is denoted as \(A_{\text{background}}(j)\). As explained earlier, the background model \((A_{\text{background}})\) is built using the background database. Thus, the score of the message \(j\) against the message \(i\) is defined as the log ratio of the two probabilities \(\left(\log_{10}(A_{\text{author}}^i/A_{\text{background}}^j)\right)\) between \(A_{\text{author}}^i\) and \(A_{\text{background}}(j)\), then normalised by the number of characters in the message group \(j\) \((N_j)\). A normalisation is necessary as the character number of each message group is different from one group to another.

In order to give a concrete example of calculating scores, suppose that the message \(j\) is a single word of ‘chatlog’, and also suppose that the character \(N\)-gram probabilities of the message \(i\) and the background database are as given in Table 3.

**Table 3**: Fabricated character bigrams and their probabilities of a hypothetical author \(i\) and a background model.

<table>
<thead>
<tr>
<th>Bigrams</th>
<th>Author model ((i))</th>
<th>Background model</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p(h</td>
<td>c))</td>
<td>0.002</td>
</tr>
<tr>
<td>(p(a</td>
<td>h))</td>
<td>0.004</td>
</tr>
<tr>
<td>(p(t</td>
<td>a))</td>
<td>0.003</td>
</tr>
<tr>
<td>(p(l</td>
<td>t))</td>
<td>0.006</td>
</tr>
<tr>
<td>(p(o</td>
<td>l))</td>
<td>0.009</td>
</tr>
<tr>
<td>(p(g</td>
<td>o))</td>
<td>0.012</td>
</tr>
</tbody>
</table>
The degree of similarity between the messages \((i\ and\ j)\) is probabilistically calculated as the product of all relevant \(N\)-gram probabilities of the message \(i\), which is \(1.5552e^{-14}\) \((= 0.002*0.004*0.003*0.006*0.009*0.012)\). On the other hand, the degree of typicality of the message \(j\) against the background model is also the product of all relevant bigram probabilities of the background model, which is \(1.62792e^{-12}\) \((= 0.005*0.009*0.008*0.017*0.019*0.014)\). The log\(_{10}\) ratio between them is \(-2.019847\) \((= \log_{10}(1.5552e^{-14}/1.62792e^{-12}))\), which needs to be normalised by the number of characters in the message \(j\) \((= 7)\), resulting in the final score of \(-0.2885495\) \((= -2.019847/7)\). In this example, bigrams were used for the sake of simplicity. The process is the same even if higher-order \(N\)-grams are used (e.g. trigrams).

The \texttt{ngram-count} and \texttt{ngram} functions of the \textit{Speech Technology and Research Laboratory Language Modelling Toolkit (SRLM)}\(^9\) are employed in this study. The former function is for building the character \(N\)-gram model of a given message and also for building the background model from the background database. The latter function is for calculating the similarity between the model of a given message (e.g. \(i\)) and the target message (e.g. \(j\)) as log probability and also the typicality between the background model and the target message. A different number (from 1 to 9) was set for the minimal count of the character \(N\)-grams. That is, all \(N\)-grams of which occurrences are less frequent than the set minimal count were discounted with a value of 0. Trigrams were used in this study.

### 3.3 Logistic Regression Calibration

As previously indicated in §3.2.1 and §3.2.2, the raw output values (scores) of the two target procedures may be calibrated before they are interpretable as LRs. Scores are in fact LRs in the sense that their values indicate the degree of similarity between two samples (e.g. messages) in comparison, having taken into account their typicality in terms of a background population (Morrison, 2013, p. 2). However, the magnitude of a score cannot be interpreted as the strength of evidence when it is not calibrated. The most commonly used technique for calibration is logistic regression (Brümmer & du Preez, 2006), which performs a monotonic shift on the scores in the logarithmic domain. The weights involved in the shift are calculated from training data for which the development

\[\text{http://www.speech.sri.com/projects/srilm/}\]
database of 38 authors was used. The *FoCal toolkit*\(^{10}\) was used for logistic-regression calibration in this study. The reader is also encouraged to read Morrison (2013) for a more detailed mathematical exposition of calibration.

### 3.4 Assessment

In this study, the log-likelihood-ratio-cost \((C_{\text{llr}})\) (Brümmer & du Preez, 2006) is used to assess the performances of the FTC systems. The \(C_{\text{llr}}\) was originally developed to assess the performance of LR-based speaker recognition systems, but it is applicable to any type of LR-based comparative systems. Traditionally, metrics based on ‘error vs. not error’, such as equal error rate (EER), have been used for assessing the performance of LR-based systems. However, EER is not an appropriate metric for the presentation of forensic evidence because it only measures the error (or success) rate of binary decision making (correct vs. incorrect) (Morrison, 2011). While EER can be a useful metric for assessing the overall discriminability of a system, it does not provide any means of assessing the gradient strength of LRs. This is the motivation behind the development of the \(C_{\text{llr}}\) metric, for which the formula is given in 4).

\[
C_{\text{llr}} = \frac{1}{2} \left( \frac{1}{N_{Hp}} \sum_{i}^{N_{Hp}} \log_2 \left( 1 + \frac{1}{LR_i} \right) + \frac{1}{N_{Hd}} \sum_{j}^{N_{Hd}} \log_2 \left( 1 + LR_j \right) \right)
\]

In 4), \(N_{Hp}\) and \(N_{Hd}\) refer to the numbers of SA and of DA comparisons, and \(LR_i\) and \(LR_j\) refer to the linear LRs derived from these SA and DA comparisons, respectively. In \(C_{\text{llr}}\), LRs which support counter-factual hypotheses are penalised in proportion to how largely the LRs deviate from unity. That is, for example, \(\log_{10} LR = -10\) for an SA comparison is more heavily penalised than \(\log_{10} LR = -5\), resulting in the former value contributing to a higher \(C_{\text{llr}}\) value than the latter. Since \(C_{\text{llr}}\) is a metric based on cost, a smaller \(C_{\text{llr}}\) indicates better performance.

It is common that the magnitudes of the derived LRs are visually represented using Tippett plots. The interpretation of Tippett plots is explained in detail in §4 in which the results of the current study are given.

\(^{10}\) https://sites.google.com/site/nikobrummer/focal
EER is also used in this study not for comparing the performance of the two procedures but for simply showing the discriminability potential of each procedure.

4. Results and Discussions

Table 4 contains the best performing results of the two procedures in terms of $C_{llr}$ and their EERs.

Table 4: Performance evaluation with $C_{llr}$. Features/Minimal count = best performing feature set for the MVKD procedure or best performing minimal count for the character N-gram procedure. EER = equal error rate.

<table>
<thead>
<tr>
<th>Procedures</th>
<th>Features/Minimal count</th>
<th>$C_{llr}$</th>
<th>EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVKD</td>
<td>2,6,11,12</td>
<td>0.604</td>
<td>20%</td>
</tr>
<tr>
<td>N-grams</td>
<td>7</td>
<td>0.807</td>
<td>25%</td>
</tr>
</tbody>
</table>

For the MVKD procedure, the best $C_{llr}$ value (= 0.607) was obtained with the four features of ‘Type-token ratio (TTR)’ (2), ‘Average character number’ (6), ‘Punctuation character ratio’ (11) and ‘Special character ratio’ (12). It is interesting to see that you don’t need all of the 12 features \(^{11}\) to get the best result, but you need as few as four features. This could be due to a ceiling effect. Another interesting point is that three features out of the four are character-based.

As for the character N-gram procedure, the system yielded the best result ($C_{llr} = 0.807$) when the minimal count was set as 7. As can be seen in Figure 2, there are a few of ups and downs in performance with the minimal count being set as 1 to 6, after which the performance stays relative stable.

\(^{11}\) The $C_{llr}$ value is 0.637 with all of the 12 features.
Comparing the $C_{dr}$ values of the two procedures, the MVKD procedure ($C_{dr} = 0.604$) outperforms the character N-gram procedure ($C_{dr} = 0.807$).

The superiority of the MVKD procedure to the character N-gram procedure is also well observed in the Tippett plots given in Figure 3. Tippett plots show the cumulative distributions of LRs ($\log_{10}LR$) separately for SA and DA comparisons. In the Tippett plots given in Figure 3, the LRs of the SA comparisons are plotted from the lowest to the highest (the black curve to the right) while the LRs of the DA comparisons are from the highest to the lowest (the grey curve to the left). In the Tippett plots, $\log_{10}LR$s are used, in which case the neutral point is $\log_{10}LR = 0$. The neutral point means that there is no support for either hypothesis. Tippett plots show how strongly the derived LRs not only support the correct hypothesis but also misleadingly support the contrary-to-fact hypothesis. An LR which is further away from $\log_{10}LR = 0$ more strongly supports either hypothesis, regardless of its truth. EER can also be observed from Tippett plots, which is the crossing point of the two curves, indicated by an arrow.
It can be observed from Figure 3b that although there are some strong consistent-with-fact LR values for both SA and DA comparisons, the strength of the LR is generally weak for the character N-gram procedure in that the large part of the curves is very close to the unity. Whereas the LR values of the MVKD procedure (refer to Figure 3a), in particular for the DA comparisons, are further away from the unity. Furthermore, the discriminability of the system is better for the MVKD procedure than the character N-gram procedure as can be seen in the ERRs which are indicated by the arrows in Figure 3; EER = 20% for the MVKD procedure and EER = 25% for the character N-gram procedure. This better discriminability also contributed to the better $C_{dir}$ value for the MVKD procedure.

It can be seen from Figure 3 that the magnitudes of the contrary-to-fact SA and DA LR values are greater for the MVKD procedure than for the character N-gram procedure. This must have negatively impacted on the $C_{dir}$ value of the MVKD procedure. However, 1) as mentioned above, the EER is better in the MVKD procedure than in the N-gram procedure; 2) the consistent-with-fact SA and DA LR values are stronger in the MVKD procedure than in the N-gram procedure; and 3) only a couple of contrary-to-fact SA and DA LRs of the MVKD procedure are greater than those of the N-gram procedure. Thus, as a whole, it can be considered that the $C_{dir}$ value of the MVKD procedure is still lower than that of the character N-gram procedure.
It was pointed out above that the best-performing feature combination for the MVKD procedure is the feature set of ‘Type-token ratio (TTR)’, ‘Average character number’, ‘Punctuation character ratio’ and ‘Special character ratio’; three out of these four features being character-based features. This point indicates that the constellation of characters making up a message carries useful authorial information for authorship analysis. However, it is interesting to see that the character-focused procedure, the character N-grams, underperforms the MVKD procedure. Some reasons can be considered for this.

First of all, the MVKD procedure uses an adaptation technique to model each group of messages with the mean vector and variance/covariance matrix; the within-group variance/covariance (the variance/covariance of each message group) are constant, and the variance/covariance are adapted from the background model. In particular, when the data is sparse, it is well-known that this type of adaptation will work well (Reynolds et al., 2000). I judge that this adaptation technique has contributed to the better performance of the MVKD procedure as the word number used for modelling each message group was as low as 500 words.

Another possible reason is that the features used in the MVKD procedure can capture the various characteristics of the authorial attributions. As reported in Ishihara (2014b), besides the character-based features, features relating to vocabulary richness (e.g. the first three features given in Table 2) are good features to be used. Thus, the ‘Type-token ratio (TTR)’, which is a feature of vocabulary richness, may have added extra information to the model, and therefore, the MVKD procedure yielded the better result than that of the character N-gram procedure.

5. Future Studies

The character N-gram technique which was employed in the current study is a standard one. However, there are various types of modelling techniques based on character N-grams (Robert Layton et al., 2010; R. Layton et al., 2012). We need to empirically test the validity of these techniques as well.

One of the strengths of the LR framework is that it allows the LRs estimated for different types of evidence to be combined (e.g. LRs from voice evidence and LRs from
DNA evidence) by multiplying them\(^{12}\) to generate single overall LR, if they are not correlated. If they are correlated, the technique called fusion needs to be employed to produce an overall LR. Theoretically speaking, if the two types of evidence\(^{13}\) are contradicting in their LRs, the overall LR would be weak, while if the two types of evidence support the same hypothesis in their LRs and they have complementary information, the fused LR will be stronger. If the two pieces of evidence share the same information, the fusion would not bring any improvements. Thus, it is interesting to see whether the fusion will improve/deteriorate the overall performance, while investigating how similar/different the LRs derived from the two different procedures are.

6. Acknowledgements

I truly appreciate the comprehensive and useful comments of the anonymous reviewer not only for enhancing the quality of this paper but also for that the comments gave a strong sense of direction to my FTC studies.

References


---

\(^{12}\) If the LR estimated from DNA evidence is 10, and that estimated from voice evidence = 100, and if they are not correlated, the combined LR should be 1000 (=10*100). In a logarithmic domain (e.g. log\(_{10}\)), the combined log\(_{10}\)LR is 3 (= log\(_{10}(10)+log_{10}(100)\)).

\(^{13}\) The number of the types of evidence is not limited to 2. Any number of types of evidence can be fused.


~ 150 ~
Appendix I. Normal QQ plots for the twelve parameters given in Table 2 with a true theoretical normal distribution (grey line). 1 = Yule’s I; 2 = TTR; 3 = Honoré’s R; 4 = Average token number per message line; 5 = SD of the token number per message line; 6 = Average character number per message line; 7 = SD of the character number per message line; 8 = Upper case character ratio; 9 = Digit character ratio; 10 = Average character number per token; 11 = Punctuation character ratio; and 12 = Special character ratio.
## Appendix 2. Correlation coefficients between the twelve features. Those values which show stronger than a moderate correlation (≥ 0.5 or ≤ -0.5) are in bold face.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>0.796</td>
<td>0.915</td>
<td>1.000</td>
<td>0.796</td>
<td>0.915</td>
<td>1.000</td>
<td>0.796</td>
<td>0.915</td>
<td>1.000</td>
<td>0.796</td>
<td>0.915</td>
</tr>
<tr>
<td>0.703</td>
<td>0.452</td>
<td>0.730</td>
<td>-0.061</td>
<td>0.083</td>
<td>0.247</td>
<td>0.083</td>
<td>0.247</td>
<td>0.083</td>
<td>0.247</td>
<td>0.083</td>
<td>0.247</td>
</tr>
<tr>
<td>0.720</td>
<td>0.585</td>
<td>0.720</td>
<td>-0.052</td>
<td>0.087</td>
<td>0.235</td>
<td>0.087</td>
<td>0.235</td>
<td>0.087</td>
<td>0.235</td>
<td>0.087</td>
<td>0.235</td>
</tr>
<tr>
<td>0.686</td>
<td>0.543</td>
<td>0.686</td>
<td>-0.072</td>
<td>0.036</td>
<td>0.211</td>
<td>0.036</td>
<td>0.211</td>
<td>0.036</td>
<td>0.211</td>
<td>0.036</td>
<td>0.211</td>
</tr>
<tr>
<td>0.081</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
</tr>
<tr>
<td>0.083</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
</tr>
<tr>
<td>0.238</td>
<td>0.111</td>
<td>0.238</td>
<td>0.111</td>
<td>0.238</td>
<td>0.111</td>
<td>0.238</td>
<td>0.111</td>
<td>0.238</td>
<td>0.111</td>
<td>0.238</td>
<td>0.111</td>
</tr>
<tr>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
<td>0.079</td>
</tr>
<tr>
<td>0.083</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
</tr>
<tr>
<td>0.083</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
</tr>
<tr>
<td>0.083</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
</tr>
<tr>
<td>0.083</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
</tr>
<tr>
<td>0.083</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
<td>0.242</td>
<td>0.154</td>
</tr>
<tr>
<td>0.083</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
<td>0.240</td>
<td>0.139</td>
</tr>
</tbody>
</table>
Loan phonology in Murrinhpatha

John Mansfield
University of Melbourne
jbmansfield@gmail.com

Abstract. Murrinhpatha has in recent decades absorbed a large number of lexical borrowings from English and Kriol. Some sound patterns in these loanwords are assimilated to native Murrinhpatha phonology, while other patterns are imported, thus extending the Murrinhpatha phonological inventory. There is also substantial variation of loan phonology, both between words and between speakers. This paper describes some salient patterns in Murrinhpatha loan phonology and explores how these combine elements of the two source phonologies.

Keywords: Phonology, Language contact, Loanwords, Loan phonology

1. Introduction
All languages have porous lexical borders, with loanwords constantly being absorbed from other languages. Often these loanwords are pronounced differently in the receiving language, assimilating to its own phonological patterns. But sometimes loanwords, even those that become frequent and widely accepted in the receiving language, maintain sound patterns that are otherwise foreign. Most brief descriptions of loan phonology treat these two scenarios as quite distinct, thereby maintaining clear limits on what should count as part of a language’s phonological system, and what should count as foreign. The bowdlerised version implies a clear distinction between assimilation (maintaining the system), and importation (updating the system). But more nuanced studies show that loan phonology produces both hybrid sound patterns and gradience, rather than a neat separation of native and foreign phonological systems.
In this paper I report on loanwords from English in Murrinhpatha, an Aboriginal language of northern Australia. Such loanwords undergo some phonological assimilation to Murrinhpatha patterns, but also import some English sound patterns that are otherwise foreign to Murrinhpatha. But importation and assimilation can occur simultaneously in different segments of the same word, or in variable renditions of the same word. Loanwords also exhibit some phonological patterns that do not match either English phonology, nor Murrinhpatha phonology, and should instead be characterised as hybrid patterns. English loanwords constitute a substantial proportion (around 20%) of lexical tokens in contemporary Murrinhpatha speech, so their pronunciation cannot be dismissed as an edge phenomena in the phonology of the contemporary language.

There has previously been very little study of loan phonology in Australian Aboriginal languages, despite the intensive lexical borrowing occurring in all surviving languages. The only dedicated studies I am aware of are a study of Macassan loanword assimilation among various languages on the north coast (Evans, 1992), and an unpublished thesis describing English loanword assimilation in a selection of Pama Nyungan languages (McManus, 2008). Both of these studies describe only the assimilation of foreign phonology, and make no significant mention of imported sound patterns. Meanwhile, more empirically detailed research on the phonology of the mixed language Gurindji Kriol does show the integration of English-based sound patterns in a generally Aboriginal phonological system (Jones & Meakins, 2013; Jones, Meakins, & Muawiyath, 2012). But such phonological mixing is certainly not limited to mixed languages: as we will see in this paper, Murrinhpatha imports sound patterns from English, which blend with indigenous sound patterns in interesting ways. It is expected that research into loan phonology in other Aboriginal languages would yield similar results.

2. Loan phonology
The foundational modern studies of loan phonology by Haugen (1950) and Weinreich (1953) focus on the assimilation of foreign material to the sound patterns of a speaker’s L1 phonology. For example, Weinreich analyses how
Schwyzerdtsch speakers in Grisons, Switzerland, mispronounce the Romansh language of the neighbouring village, arguing that a Romansh palatal stop /c/ is reinterpreted by Sch. speakers as their L1 affricate /tʃ/; or again, the Romansh geminate consonant /ss/ is conflated by Schwyzerdtsch speakers with the single consonant /s/ (p. 17). On the other hand, there are well known cases of historical phonology involving importation of sound patterns. For example, though /v/-initial words were once absent from English, and earlier Latin borrowings were nativised to /f/ (e.g. *Fergilius* < Lat. *Virgilius*), the weight of French borrowings in Middle English eventually retained the phoneme, so that contrasts such as *feel* vs *veal* (< Fr. *veel*) could be sustained (Labov, 1994, p. 332; Lass, 1992, p. 58).

In the case of English, the absorption of Romance and Greek words into a genetically Germanic language means that the lexicon can be analysed as having distinct lexical strata – not just as a point of etymological history, but as a generalization about the sound patterns shared by sets of words, especially with respect to derivational processes and associated stress patterns. In classic generativist accounts (Chomsky & Halle, 1968, pp. 171–174; Lightner, 1972), it was proposed that such *lexico-phonological strata* are encoded as diacritics in the lexicon: a class of lexical items are marked as belonging to a stratum, which makes them subject to special rules in generating surface forms. But subsequent studies problematised the neat generalisations of lexico-phonological strata. Analysis of French and English loanwords in Russian showed that they do not form neat strata, because some loanwords simultaneously exhibit “Russian” and “foreign” patterns (Holden, 1976). For example *kôfe* “coffee” imports an /f/ segment that is absent in native Russian phonology, but palatalises this consonant under the influence of the following vowel, thus also showing some assimilation to a native pattern. The range of foreign structures in these loanwords appear to form a graduated hierarchy from the most imported to the most assimilated, which Holden takes as evidence for a diachronic integration process assimilating some structures more rapidly than others.

Itô and Mester (1995) provide a nuanced analysis of how Chinese and European loanwords coexist with an earlier indigenous “Yamato” stratum in the Japanese lexicon. They show that the lexicon is better modelled by overlapping and cross-
cutting sound patterns, than by lexical subsets. For example /p/ segments always occur as part of heavy, homorganic /pp/ or /mp/ clusters throughout the Yamato (indigenous Japanese) lexicon; but this pattern also characterises Chinese loanwords, but not mimetic words or European loanwords. Cross-cutting this is a progressive voicing assimilation pattern in nasal-stop clusters, which characterises Yamato and mimetic words, but not Chinese or European loanwords (Figure 1).

The various patterns of this type model the lexicon as a multiply overlapping Venn diagram, with a “core” area occupied mostly by Yamato words, surrounded by various other intermediate intersections, and a peripheral area in which the least restricted patterns occur, though even these share some sound patterns with the rest of the lexicon. European loanwords occupy the peripheral area, but there are also loanwords that don’t follow the same patterns as other words from the same etymological source, supporting the analysis of phonologically characterised strata, rather than source-language generalisations.

Itô and Mester model the lexicon as cross-cutting phonological patterns, with each intersection populated by a subset of lexical items. But they also note that these intersections are something of an abstraction, glossing over the variable
phonological form of words (p. 821). Indeed, variationist research on loan phonology shows that words show substantial variation with respect to their membership in lexico-phonological strata. A large-scale, longitudinal study of English borrowings in Canadian French (Poplack, Sankoff, & Miller, 1988; Poplack & Dion, 2012) shows that some borrowings are quite consistently assimilated, while others only variably so. Probability of assimilation is influenced by both the word, and the speaker: it is more likely to be assimilated if it is a frequently used word, and if it is spoken by someone with low competence in the source language. A corollary of this is that borrowings are more likely to be assimilated phonologically if they are diffused among more speakers (Poplack et al., 1988, pp. 73–74). Further social influences on variable loan phonology have been shown for Spanish loanwords in Mexicano (Uto-Aztecan). These loanwords are more likely to import Spanish phonology if they refer to the domains of technology, education and commerce, in which Spanish has a more prestigious position. Importing of Spanish phonology was also more likely when the interlocutor was proficient in Spanish (Lev-Ari, San Giacomo, & Peperkamp, 2014, pp. 669–673).

The overall effect of sociolinguistic variation is that speakers move lexical items in and out of lexico-phonological strata according to their goals in social interaction. Every time I order a croissant, I make a decision about what sort of message and persona I want to project, and to whom, and accordingly I select between [kwasã ~ kwasɔnt ~ krəzɔnt], and gradient versions in between (Matras, 2009, p. 110). Inasmuch as my day-to-day language use contains a high number of loanwords, its sound patterns therefore reflect not a static inventory of phonemes and word-formation processes, but a rather fluid intersection of possibilities. Similar sociolinguistic variation occurs in Murrinhpatha loan phonology, and indeed is consciously referenced by speakers as an index of relationships to whitefellas, to settlement and modernisation. For example, a group of young men told me how their grandparents, whom they characterised as kardu putj-mawu “from the bush”, used to say tjitj o’clock, whereas they, the new generation, say, six o’clock. A full investigation of such variables would be of great interest, but is beyond the scope of this paper, in which I limit myself to some comments on differences between careful and spontaneous speech, and in general try to abstract away from stylistic...
variation to focus on the most common loan phonology variants produced in the vernacular speech of young men.

3. **English loanwords in Murrinhpatha**

Murrinhpatha offers an excellent illustration of loan phonology phenomena, since it is absorbing a large wave of loanwords from English, and the phonologies of these two languages are so different that there are many points on which we can observe whether a sound pattern has been imported or assimilated. I provide here a selective description of loan phonology in Murrinhpatha, highlighting phenomena that demonstrate the mixing of sound patterns from the two languages. I do not attempt to comprehensively describe all loan phonology phenomena in Murrinhpatha (but for more details see Mansfield, 2014), but instead focus on three segmental structures that are salient examples of English phonology foreign to Murrinhpatha: fricatives, diphthongs, and syllable onset clusters.

3.1 **Sociolinguistic setting**

Murrinhpatha is one of the few Australian Aboriginal languages that remains vibrant, with children still raised as L1 speakers, and a growing speaker base of some 3000 people in the remote town of Wadeye, Northern Territory, and surrounding smaller settlements. One reason why the speaker base is growing is that various other language groups (aka “tribes”) from the area have shifted from their heritage languages to Murrinhpatha; a few older people remain bilingual in other Aboriginal languages, but young people in Wadeye generally acquire only Murrinhpatha as L1. English, and the English-lexified Kriol that is common across northern Australia, are used as lingua francas for communication with non-Aboriginal people, and Aboriginal people from other regions, respectively.

The role of English in Wadeye is ubiquitous, but somewhat shallow. Non-Aboriginal people constitute only a small minority of the population, but dominate most of the management and service jobs in the town; therefore Murrinhpatha speakers must use English frequently, but mostly for a limited range of transactional tasks, like buying groceries. The role of Kriol is quite
different, as it is used for stronger ties, usually kinship by marriage, with Aboriginal people from elsewhere who are staying in Wadeye. English and Kriol have largely cognate vocabularies, distinguished by phonological form; but because the assimilatory processes in Kriol are similar to those in Murrinhpatha, it is usually impossible to determine whether loanwords in Murrinhpatha have been absorbed from English or from Kriol. It is likely that many loanwords are encountered by Murrinhpatha speakers in both their English and Kriol forms, e.g. both outside and atsait, both sell and tjelim. In this paper I refer to “English” loanwords as a shorthand, keeping in mind that most words are probably being absorbed simultaneously from both English and Kriol.

Young Murrinhpatha speakers’ confidence and competence in English is highly variable, ranging from rudimentary to quite fluent. Young men tend to have quite limited fluency, while women are rather more confident. Formal education in Wadeye is bilingual, though English takes the dominant role; however school attendance is very low, especially among boys. Nonetheless, all young people make prolific use of English loanwords – and given that these words are used both by skilled bilinguals and by those who are almost monolingual in Murrinhpatha, they must be considered borrowings and not codeswitches (Matras, 2009, p. 114). As an illustration of the scope of borrowing, I counted tokens in a recorded natural conversation sequence of about 15 minutes’ duration, between three young men who have limited English bilingualism. Borrowings account for 22% (58/265) of content words in this sequence. The 58 borrowed tokens comprise 47 distinct borrowed lexemes, including back road, shortcut, supper, finish, taste, tea, mob, family, story, money, pocket, save, contract, pay day, ten o’clock, two-hundred, killing time, really, lump, friend, policeman, late, heavy, same. These include a mixture of cultural/technological borrowings (e.g. tea, pocket) and gratuitous borrowings (e.g. family, killing time).

3.2 Data and method

The data used here is all drawn from the speech of young men age 18–40, who constitute an emic age grade kardu kigay (person + youth). Since 2011 I have spent substantial time doing linguistic and ethnographic fieldwork in Wadeye, working mostly with kigay. The data used here is picture-elicited speech data from 25 kigay, and more naturalistic conversational recordings of 9 kigay, most of whom
also appear in the elicited speech. The elicited speakers each provided names or
descriptions for 108 pictures, of which about a dozen pictures were technological
items (hammer, bed) that predictably elicited loanwords, while other items (river,
man in pain) sometimes produced gratuitous loanwords, despite the existence of
indigenous Murrinhpatha alternatives. Each speaker therefore produced 15–25
loanwords.

It is important to recognise the limitations of elicited speech in studying loan
phonology. Elicitation is likely to make speakers highly conscious of how they are
speaking, and this may lead to them to speak in ways that are hardly
representative of their main day-to-day speech practices (Labov, 1972). For loan
phonology this presents the risk that speakers make a special effort to pronounce
loanwords “properly”, and therefore reduce the degree of assimilation they would
produce in more naturalistic speech. To address this risk, I treat all phonological
imports in the elicited speech with suspicion, and only claim them as genuine
imports where I have identified additional data from the conversational
recordings, where the same import is in evidence.

Variationist analysis of Murrinhpatha loan phonology would be of great interest,
but the data drawn upon here does not facilitate such an analysis. The picture-
elicitiation method produced similar wordsets from the 25 participants, but with
each speaker using slightly different words and phrases to describe the pictures.
Given the fairly small number of loanword tokens elicited, the data does not
provide enough comparable data for statistical analysis. Therefore this study does
not provide a quantitative analysis of variation, but rather an overview of the
range of sound patterns produced by young Murrinhpatha men.

At various points in this study I refer to “old borrowings” or “recent
borrowings”, citing older borrowing as a likely explanation for why some lexical
items are more assimilated than others (Hill & Hill, 1986, p. 198; Poplack et al.,
1988, p. 85). There is a risk of circularity here (i.e. presuming that a borrowing is
old because it is unusually assimilated), however I only posit a borrowing as “old”
where there are historical reasons to do so. An “old borrowing”, then, is one that
is likely to have already been widespread during the Catholic mission era (1939–
1975) when the Murrinhpatha speech community was first exposed to substantial
English contact, as opposed to “recent borrowings” of the post-missionary era. Clear examples of old borrowings are *tina* ("dinner") and *fjaba* ("supper"), meal-times that were established as part of the mission routine (Catholic Mission Port Keats, 1969), while *mabal* ("mobile phone") and *ailatj* ("(Toyota) Hilux") are clear examples of recent borrowings. However the suggestion that unusually assimilated lexical items can be explained by age of borrowing is only a hypothesis, as an alternative historical explanation that generally cannot be ruled out is that the item may have been borrowed indirectly via another Aboriginal language that assimilates more restrictively than Murrinhpatha.

### 3.3 Fricatives

English has obstruent manner contrasts between stops, affricates and fricatives, while Murrinhpatha (and Aboriginal Australian phonologies in general) does not have direct stop/fricative contrasts. However Murrinhpatha does have two series of obstruents, which are primarily distinguished by voicing, although they also involve differences of constriction type and timing (Butcher, 2004; Mansfield, 2014, pp. 141–143). Most obstruents of both series are produced canonically as stops, but the voiced dental and velar obstruents are almost always fricated, while palatals are affricated in syllable onsets and stopped in syllable codas. However, crucially for the argument below, Murrinhpatha obstruent contrasts are maintained only in syllable onsets, and not in codas, where all obstruents are unreleased voiceless stops.1 We can characterise this asymmetry more generally as a greater sensitivity to obstruent timing/constriction in onsets, as opposed to codas, and we will see that this generalisation holds good with regard to new obstruents imported from English.

The consonant inventory of Murrinhpatha, prior to English loanwords, is illustrated in Figure 2.

---

1 There is a tradition of analysing Australian phonotactics in terms of “word templates” rather than syllables, based on the argument that purportedly disyllabic words would have such different constraints on their first and second “syllables” that this unit of analysis is of limited value and a whole-word analysis is to be preferred (Baker, 2014, pp. 143–145; Dixon, 1980, pp. 159–178; Hamilton, 1996). However in Murrinhpatha there are rather fewer differences between first and second “syllables”, so that this prosodic unit does have descriptive utility (Mansfield, 2014, pp. 152–3). Nonetheless, a full analysis of arguments for and against the syllable in Murrinhpatha has never been undertaken.
English loanwords import three voiceless fricatives /f, s, ʃ/, while the voiceless affricate /tʃ/ is imported in careful speech, but probably not in natural speech. The English voiceless fricatives /θ/ and /h/ are assimilated to /t̪/ and to zero respectively, and in general voiced fricatives are merged with their voiceless counterparts. /f, s/ are imported only in syllable onsets, while in codas they assimilate /f/ → /p/ and /s/ → /c/ ~ t/. On the other hand /ʃ/ is imported both in onsets and codas in almost all instances, which includes a number of natural-speech tokens for bush, rush and finish. I would have predicted /ʃ/ to pattern with the other imported fricatives /f/ and /s/ in this context (i.e. by assimilating to word-final /c/), and I have no ready explanation for why it is borrowed differently.

Lexemes that consistently import /f, s, ʃ/ give rise to the following contrasts.

<table>
<thead>
<tr>
<th>PERIPHERAL</th>
<th>LAMINAL</th>
<th>APICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>Dental</td>
<td>Alveolar</td>
</tr>
<tr>
<td>Voiceless obstruent</td>
<td>p</td>
<td>t̪</td>
</tr>
<tr>
<td>Voiced obstruent</td>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>η</td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td>l̪</td>
</tr>
<tr>
<td>Trill / flap</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>j</td>
</tr>
</tbody>
</table>

*Figure 2. Murrinh-Patha consonant phonemes*
Fricative versus stop:

fritʃ “fridge”
pritʃ “bridge”
finish “finish”
pirnipirn “armband” (MP indigenous lexeme)

Fricative versus affricate:

seben “seven”
Tjebin (language name)
shit “shit”
tjitj “bacteria”
she “share”
tje “ear”

Alveolar versus palatal fricative:

seip “save”
klinʃheip “clean-shaved”

Alveolar fricative versus dental stop:

singten “sink (down)”
thingk /t̪iŋk/ “think”

The integration of fricatives and affricates in Murrinhpatha is very similar to that reported for Roper Kriol (Baker, Bundgaard-Nielsen, & Graetzer, 2014); this may reflect the influence of Kriol varieties in the borrowing of originally English lexemes in Murrinhpatha, or it may simply reflect common assimilation strategies used by Aboriginal bilinguals, whose phonological inventories tend to be very similar (Evans, 1995). In Areyonga Pitjantjatjara the integration of fricatives is again very similar, except that in addition the /s, z/ voicing distinction is claimed to be imported (Langlois, 2004, pp. 41–43).

Most Murrinhpatha loanwords with initial /ʃ, s, ʃ/ quite consistently either import the fricative or assimilate it to the indigenous stop or affricate. But a few words vary greatly between speakers, with no pronunciation showing clear dominance. The picture-elicited data gives an overall view of the scope of variation, since some borrowed lexemes have tokens from 10–20 kigay. This data shows the lexical specificity of the pattern; for example 12 speakers unanimously import friding → /ʃɹɪd/, while 9 speakers unanimously assimilate football → /puʈpuɭ/. An

---

2 The only difference I find from Baker et al’s brief report on fricatives is that they report English dental fricatives being assimilated to Kriol palatal affricates, which does not occur in Murrinhpatha loan phonology.
example of a “mixed” lexeme is *fishing*, for which a handful of speakers import /f/, the majority assimilate it to /p/, and two speakers produce both imported and assimilated tokens consecutively.

The variable production of stop/fricative and affricate/fricative contrasts belongs to a lexico-phonological stratum that is not purely either Murrinhpatha or English in its sound patterns. The use of contrasting obstruent manners is derived from the English pattern, but it is not identical to English, in that it mostly applies in syllable onsets, but not in codas. In this sense it is a hybrid pattern, drawing on English phonology but filtered by some other influence. We noted above that Murrinhpatha has obstruent voicing contrasts, but these are collapsed in syllable codas; arguably the lack of stop/fricative coda distinctions is influenced by a pre-existing pattern of using only unreleased voiceless stops in syllable codas. This sort of constraint seems to have a basis in universal articulatory or perceptual mechanisms, since there is strong cross-linguistic evidence for greater manner constraints in codas (Blevins, 1995; Parker, 2001).

### 3.4 Diphthongs

Murrinhpatha has just four vowels, /a, e, i, u/, which are realized approximately as [ɐ, ɛ, ɪ, ʊ]. There is also a glide consonant /j/ that can appear as a syllable coda after back vowels (e.g. *kigay* “boy”, *mardinybuy* “girl”), producing phonetic offglides [aj, uj]; however the labio-velar glide /w/ does not appear in syllable codas.

Australian English has a much larger vowel system with 12 monophthongs (counting schwa) and 5 diphthongs (Harrington, Cox, & Evans, 1997, p. 178). Most English lax monophthongs integrate transparently to similar Murrinhpatha vowels, but /æ, ə/ lack Murrinhpatha equivalents. Of the latter two, /æ/ assimilates to /e/ (e.g. *bag* → /bek/), while /ə/ is usually imported (e.g. *involve* → /inbol/). Tense monophthongs are generally assimilated to their lax counterparts, though /u:/ may be a marginal exception.3

3 The open monosyllabic realisation of /ʉː/ is imported faithfully in all instances I have checked, i.e. *two* → [tuː], *glue* → [gluː], *you* → [yuː]. Impressionistically this vowel appears to have a distinct quality from the Murrinhpatha /u/, which in open monosyllables is lengthened, but in quality remains similar to the shorter
It is diphthongs in which the most extensive importation of sound patterns occurs. There are five diphthongs that feature in English loanwords: /æi, ae, ɔɪ, æɔ, æə/. Those with a front offglide [ɪ] or [ɛ] are usually imported (e.g. paper → /peipa/), while those with a back offglide [u] or [ɔ] are usually assimilated (e.g. float → /flɔt/), essentially by losing their offglides. There is however variation among both types, such that front diphthongs /æi, ae, ɔɪ/ are usually imported in the loanword stratum, while back diphthongs /əu, æɔ/ are usually assimilated to monophthongs, at least in unmonitored speech.

The diphthongs /æi, ae/ are both imported quite consistently into Murrinhpatha as new vowel phonemes, represented here as /ei/ and /ai/ respectively. Thus tape → /teip/, make up → /meikap/, design → /disain/, knife → /naip/. I know of just a few cases of assimilation for /æe/ — the probably old borrowing bicycle → /пucιŋkul/, and the much newer, but extremely common mobile (phone), which is variably /mobul ~ mabul ~ mabal/. The latter is an interesting case in that it has undergone the sort of vowel assimilations normally associated with early borrowings, even though it is clearly a recent borrowing.

The /ɔɪ/ diphthong is more variable in its importation/assimilation. When it does assimilate, it is to a simple /u/, as in toilet → /tulet/. The borrowing of /ɔɪ/ in open syllables assimilates to the indigenous /Cuj/ syllable pattern, as in schoolboy → /kulbuy/; similarly the borrowing of /æe/ in open syllables can be regarded as a transparent integration to /Caj/, with minimal phonetic modification required, as in try → /tɻaj/.

The importation of front-offglide diphthongs, but not back-offglide diphthongs, appears to be facilitated by the glides already present in Murrinhpatha as version, i.e. /dʊː/ “cry”. This does suggest then a phonemic contrast, and a new vowel introduced to Murrinhpatha, but one which appears only in open monosyllables, and has only one really common lexical exponent, namely two.

4 I set aside here the “falling diphthongs” /ɪə, ɛə, ʊə/, since in Australian English they are often produced as monophthongs (weird → [wi:d]) or dissyllables (fear → [fiːə]) (Harrington, Cox, & Evans, 1997, p. 174); neither do I consider the /iə/ sequence as in cute, which is only treated as a single segment by some linguists (Ladefoged & Johnson, 2011, p. 93), and in any has very scarce borrowing evidence in Murrinhpatha.

5 For Roper Kriol, Baker et al (2013) report the more faithful import mobailfoun, though it is not clear whether this form appears in both careful and natural speech.
consonantal syllable codas. As mentioned above, Murrinhpatha has a front-offglide coda /j/, but not a back-offglide coda /w/. Therefore the phonetic sequences [ei] and [ui] are indigenous to Murrinhpatha (e.g. kigay [kiɣɐi]) but in terms of phonological structure they constitute entire /VC/ syllable rimes, rather than complex nuclei to which codas can be added. In English loanwords however a coda consonant can be added: save $\rightarrow$ /seip/, cane toad $\rightarrow$ /kein.tot/, time $\rightarrow$ /taim/. The syllable structure of these borrowings is here interpreted as introducing complex vocalic nuclei /ei, ai/, though they could also be interpreted as introducing a new type of coda cluster /jp, jn, jm/. However if this were a new coda cluster type, it would show no obvious affinity with the rather restricted set of coda clusters otherwise appearing in Murrinhpatha (always liquid + velar, see below), and for this reason I interpret the new sequences as complex nuclei, i.e. diphthongs.

It appears that front diphthongs have been imported much more than back diphthongs because they have indigenous phonetic models for their front offglides. In this sense the emergent diphthongs draw on sound patterns from both English and Murrinhpatha.

### 3.5 Consonant clusters

The third major segmental structure of Murrinhpatha loan phonology is intra-syllabic consonant clusters, where English presents a considerable range of novel clusters in both syllable onsets and codas. With respect to fricatives and diphthongs, we have seen that the new patterns in Murrinhpatha loan phonology are hybrid patterns combining English and Murrinhpatha elements. With respect to consonant clusters, however, the new patterns run counter to Murrinhpatha’s earlier cluster patterns. Prior to English loanwords, Murrinhpatha had syllable shapes CV, CVC and occasionally CVCC – i.e. with a few coda clusters, but no onset clusters, which is typologically unusual in global terms (Donohue, Hetherington, McElvenny, & Dawson, 2013), though it is found in other languages of northern Australia (Fletcher & Butcher, 2014, p. 111). But in loanwords, onset clusters are imported, while coda clusters are assimilated in all cases except those that match the coda clusters already allowed in Murrinhpatha. Loanword integration therefore reverses Murrinhpatha’s previous asymmetry.
towards greater coda complexity, instead following the more typologically common asymmetry towards greater onset complexity.

The following English onset clusters are imported into Murrinhpatha (with some stop voicing recategorisations, not discussed here):

<table>
<thead>
<tr>
<th>Onset Cluster</th>
<th>English Word</th>
<th>Murrinhpatha</th>
</tr>
</thead>
<tbody>
<tr>
<td>/CL/</td>
<td>plate</td>
<td>/pleit/</td>
</tr>
<tr>
<td></td>
<td>clap</td>
<td>/klep/</td>
</tr>
<tr>
<td></td>
<td>glass</td>
<td>/klac/</td>
</tr>
<tr>
<td></td>
<td>bridge</td>
<td>/pɕic/</td>
</tr>
<tr>
<td></td>
<td>truck</td>
<td>/tɻak/</td>
</tr>
<tr>
<td></td>
<td>freaking out</td>
<td>/fɻikinet/</td>
</tr>
<tr>
<td>/Cw/</td>
<td>twenty</td>
<td>/tweni/</td>
</tr>
<tr>
<td></td>
<td>quicksand</td>
<td>/kwiksen/</td>
</tr>
<tr>
<td>/sC/</td>
<td>scary</td>
<td>/skɛɻiwan/</td>
</tr>
<tr>
<td></td>
<td>speedy</td>
<td>/spidi/</td>
</tr>
<tr>
<td></td>
<td>stuff</td>
<td>/stap/</td>
</tr>
<tr>
<td>/sN/</td>
<td>sniper</td>
<td>/snaipa/</td>
</tr>
<tr>
<td></td>
<td>smart-arse</td>
<td>/spadas/</td>
</tr>
<tr>
<td>/sCL/</td>
<td>spray</td>
<td>/sprei/</td>
</tr>
</tbody>
</table>

The importation of these onset clusters is almost categorical; I know of just four lexical items that defy the trend:

- flour → /lawam/
- bloomers → /pulemic/
- schoolboy → /kulbuj/
- spoon → /pun/

*Flour* → /lawam/, in which the final /m/ is quite mysterious, is certainly among the oldest borrowings, being one of the core cultural artefacts of the early Catholic

---

6 Some English voiced stops are imported as voiceless in Murrinhpatha, both in clusters and in simple onsets. However the phenomenon is quite inconsistent among lexical items, and appears to be an area of substantial inter-speaker variation. A worthy description of voicing distinctions in Murrinhpatha loan phonology requires considerable study of phonetic variation, and is therefore beyond the scope of this paper, though an overview of the problem can be found in Mansfield (2014, pp. 216–220).

7 This is used as an adjective, meaning “very stoned”.

8 I have identified only this one token of /sm/ onset integration (JL, 2013-06-22). Further data would be required to show whether the de-nasalisation /sm/ → /sp/ is a regular pattern.
mission (Pye, 1972). *Bloomers* → /pulemic/ must be an early borrowing, since the term has not been in common use in Australian English for the last half-century. *Schoolboy* → /kulbuy/ is also likely to be an early borrowing, since schooling for all children was another of the mission’s first priorities. These exceptions suggest that there may have been a brief period at the beginning of intensive English contact when (some) onset clusters were not imported – though the range of well-established borrowings in which clusters are maintained shows that this can not have lasted very long.

Meanwhile most coda clusters are assimilated. Murrinhpatha has only a very limited range of coda clusters, all with a liquid followed by a velar consonant, e.g. /lk, ɭk, rk, lŋ/. English has a much wider range of patterns selecting from liquid/nasal + obstruent + s/z. These assimilate to Murrinhpatha simple codas by maintaining only the first segment of the cluster:

friend → /fɻen/
taste → /teis/, fast → /faswan/
cards → /kat/
Tarax⁹ → /teɻek/

The only new type of coda cluster imported is homorganic nasal + stop, and perhaps only voiceless stop in view of the assimilated /fɻen/ example above. There are just a few examples of this loan pattern, sourced either in English homorganic nasal + stop, homorganic nasal + fricative, or homorganic nasal + stop +fricative:

think → /tıŋk/
pump → /pamp/
sense → /sent/
underpants → /andapent/

These coda cluster imports differ from the assimilated codas above in that they only require a simple oral gesture (Browman & Goldstein, 1992), with the “complexity” of the coda consisting of a different timing of glottis and velum, compared to a simple nasal coda. In a simple nasal coda, the velum remains open

⁹ Tarax is an Australian drinks brand that was sold at the Wadeye shop long enough to be borrowed as the word for fizzy sweet drinks in general, *kura terrik*. The brand is no longer sold in Wadeye, but the borrowed word persists.
as long as the glottis does, while in the complex coda the velum closes before the
glottis.\textsuperscript{10} Therefore it is \textit{complex oral coda gestures} in particular that are simplified in
the coda, while complex gestures of various types are faithfully imported in onset
clusters.

The intra-syllabic consonant cluster patterns found in Murrinhpatha loan
phonology again imports some dimensions of the English pattern, but assimilates
others. However in this case the divergence from the source pattern is more
plausibly motivated by dispreference for complex codas, which is a cross-
linguistic tendency, but not a pattern previously exhibited by receiving language.
This begs the question, which remains unanswered here, of why Murrinhpatha
should have had a typologically unusual asymmetry towards complex codas in its
earlier vocabulary, but switched to greater onset complexity in its loan phonology.

4. Conclusion

Contemporary Murrinhpatha phonology does not have a unitary, consistent set of
sound patterns, but instead exhibits a mixture of Murrinhpatha-derived and
English-derived patterns. The different patterns cannot be divided neatly into
etymological subsets, just as they cannot for Russian, Japanese or Mexicano
(Holden, 1976; Itô & Mester, 1995; Lev-Ari et al., 2014), and furthermore there
are patterns that show hybrid qualities combining English and Murrinhpatha
sources. A fricative/stop contrast is imported in some lexical items, but mostly
just in syllable onsets, which is consistent with Murrinhpatha’s pre-existing system
of obstruent contrasts in onsets but not codas. Diphthongs are the second major
segmental type imported, but again only parts of the English diphthong inventory
are imported, apparently under the influence of phonetic offglides already present
in Murrinhpatha. A wide range of onset consonant clusters are also imported, but
this instance is more difficult to explain in terms of pre-existing Murrinhpatha

\textsuperscript{10} The new coda pattern does not involve an oral stop release, but rather a sharp cessation of the nasal
resonance or indeed any other airflow, contrasting with the gradually declining resonance of word ending
in a simple nasal. These codas could therefore be represented as ending either in unreleased oral stops
\textit{/ŋk, mp, nt/}, or in glottal stops \textit{/ŋʔ, mʔ, nʔ/}. 

~ 169 ~
sound patterns, which permit consonant clusters just in codas and not in onsets. In all three of these patterns, however, Murrinhpatha loan phonology shows that intense contact with a very different phonological system can have a rapid and substantial effect on the sound patterns used in a language.

References


Abstract. We investigate how wh-in-situ languages disambiguate the scope of embedded wh-phrases with prosody. Following current research on the syntax-prosody interface, we note this depends on a principle of prosodic contiguity (Richards 2011) which requires wh-phrases and the head of their associated CP to be contained in one Intonational Phrase. This alters the prosody of any intervening material. From our own research, we show Telugu resolves wh-scope ambiguity by creating a prosodic domain with a high flat pitch between the wh-word and either the embedded or matrix C where it takes scope. Khmer also makes use of prosody to disambiguate, though in a different manner, which we argue raises issues for analyses making use of principles similar to prosodic contiguity.

Keywords: Wh-in-situ, wh-movement, prosody, Telugu, Khmer

1. Introduction
Languages differ in whether they form questions by moving question words (henceforth wh-phrases), typically to the beginning of the sentence, or by leaving them in-situ. This paper presents analysis of a selection of languages which, unlike English, form questions without wh-movement.
Aside from *wh*-movement, (1b) and (2b) are equivalent: both are interpreted as interrogative sentences. A long history of research in the generative tradition seeks to answer the question implied by this contrast: why do some languages move *wh*-phrases while others do not?

In established theory, *wh*-movement is understood to take place because a structure originating in one position in a syntactic derivation bears a *wh*-feature which requires checking or licensing in some other position. This manner of approach predates Chomsky’s Minimalist Program (1995), and continues to be the foundation on which current research in this framework is built. For example, Richards (1997) proposes that *wh*-movement languages have ‘strong’ *wh*-features which cause overt movement, while *wh*-in-situ languages have ‘weak’ *wh*-features which do not. This analysis is rooted in the proposal that *wh*-movement is universal, but that *wh*-in-situ languages differ from *wh*-movement languages by moving *wh*-phrases covertly (Aoun, Hornstein, & Sportiche 1981, Huang 1982). That is, in *wh*-in-situ languages, sentences such as (1b) ultimately have an unpronounced logical form corresponding to the overt syntax of (2b). Syntactic structures corresponding to the overt and covert structures that would be proposed for Khmer interrogatives are presented in full below:
(3)  

a. Surface (overt) structure of a Khmer interrogative

```
CP
   C'
      C
         TP
            DP
                Rany
                 T
                     VP
                        V
                            DP
                                rian
```

b. LF of a Khmer interrogative, equivalent to English surface structure

```
CP
   C'
      C
         TP
            DP
                Rany
                 T
                     VP
                        V
                            DP
                                rian
```

By contrast, others argue that that some other operation allows wh-phrases to be interpreted in-situ (Aoun & Li 1991, Reinhart, 1998). Our study departs from this type of analysis in considering primarily how the scope of a wh-phrase can be expressed by means other than linear word order. Indeed, we make no claim to bear on the debate over overt vs. covert movement, as our analysis does not hinge on whether the prosody associated with wh-in-situ structures is an overt correlate of covert movement, or whether it obviates the need for such movement in the first place. We begin by noting that movement can influence interpretation in embedded contexts. Consider how, in wh-movement languages, wh-scope is

---

1 For the tree structure in 0 to faithfully represent English as well as Khmer, it would be necessary to show movement of the auxiliary from T to C to derive English subject-auxiliary inversion. We make no further reference to this distinction between the languages. Khmer has no auxiliary in this context and, moreover, subject-auxiliary inversion is not otherwise relevant to the study here.
expressed by linear position - in (4) the \(wh\)-phrase corresponding to the embedded object moves to either the periphery of the embedded clause to derive an indirect question (4a), or of the matrix clause to derive a direct question (4b).

\[(4) \quad \begin{align*}
a. & \quad \text{Lionel knows [which book] Bertie read } \text{which book} \\
b. & \quad [\text{Which book}] \text{ does Lionel know Bertie read } \text{which book} ?
\end{align*}\]

In \(wh\)-in-situ languages, given the lack of overt movement, a theoretical issue arises in that, unlike in (4), there is no linear cue as to the \(wh\)-phrase’s scope. (5) shows the persistent ambiguity that our study addresses.

\[(5) \quad \begin{align*}
\text{Boran} & \quad \text{duŋ-tʰaa Rany rian ai} \\
\text{Boran} & \quad \text{knows Rany learns what}
\end{align*}\]

\textit{Khmer}

\begin{align*}
\text{Matrix } \text{\(wh\)}\text{-scope:} & \quad \text{‘What does Boran know Rany is learning?’} \\
\text{Embedded } \text{\(wh\)}\text{-scope:} & \quad \text{‘Boran knows what Rany is learning’}
\end{align*}

We show in the remainder of this paper how prosody serves to disambiguate \(wh\)-scope in \(wh\)-in-situ languages. We will additionally identify a conceptually problematic case for these theories, which we argue to be attested in our own research, but argue that this is best addressed with the issue of \(wh\)-scope disambiguation in mind.
2. Movement and prosody

In redress to the abstract nature of the research cited above, we situate our research in a developing movement in theoretical linguistics which is concerned with identifying concrete correlates of different movement dependencies in a language’s phonology and prosody.

2.1 The syntax-prosody mapping

It has been observed that in \textit{wh}-in-situ languages there is often a unique prosodic structure or intonation associated with the \textit{wh}-phrase (Ishihara 2003, Richards 2010). It is argued that this prosody is reflective of a particular prosodic structure which deviates from the language’s regular prosodic constituency, but in a systematic and rule-governed way. Before this observation can be developed, though, it is necessary to present the framework we take to derive the mapping from syntactic to prosodic structure.

The principles governing the syntax-prosody mapping are designed to create a linear, contiguous prosodic structure from a hierarchical and recursive syntactic structure (Selkirk 1984, Truckenbrodt 1995). In a simplified manner, we follow these principles in deriving a prosodic structure, consisting of a series of Intonational Phrases (IPs), by identifying either the left or right edge of a language’s syntactic phrases as points of division between IPs, or as \textit{prosodic boundaries}. To illustrate, in the structure below the left edge of each DP marks a boundary between IPs. Thus, each DP \textit{begins} a new IP.

\begin{align*}
\text{Syntax} & :\quad [TP \ [DP \ D \ NP ] \ [VP \ [DP \ D \ NP ] \ V ]] \\
\text{Prosody} & :\quad (IP \ D \ NP ) \quad (IP \ D \ NP \ V )
\end{align*}

As will be seen, the ability to create a boundary between IPs is not just a property of the edges of DPs. The edges of any phrase, or \textit{XP}, can do so. Thus, if the left...
edges of the NPs above also mapped onto a prosodic boundary, the resulting prosodic structure would consist of four IPs:

\[
(7) \begin{align*}
    \text{Syntax} & : \quad [\text{TP} \ [\text{DP} \ D \ [\text{NP} \ N]] \ [\text{VP} \ [\text{DP} \ D \ [\text{NP} \ N] \ V ])] \\
    \text{Prosody} & : \quad (\text{IP} \ D) \ (\text{IP} \ N) \ (\text{IP} \ D) \ (\text{IP} \ N \ V )
\end{align*}
\]

With an outline of the syntax-prosody mapping in place, we can begin analysis of the prosody associated with \textit{wh}-phrases.

\subsection*{2.2 Prosodic Contiguity}

The noted prosody associated with \textit{wh}-phrases is commonly proposed to surface in order to satisfy a principle requiring the structure in question to be placed into a single IP with an associated head in the relevant CP layer of the derivation (Hirotani 1996). We adopt Richards' (2011) principle of \textit{prosodic contiguity}.

\[(8) \quad \textit{Prosodic Contiguity} \quad (\text{Richards} \ 2011)\]

Given a \textit{wh}-phrase \(\alpha\) and a complementizer \(C^0\) where \(\alpha\) takes scope, \(\alpha\) and \(C^0\) must be separated by as few Minor Phrase boundaries as possible, for some level of minor phrasing.

Richards proposes \textit{Prosodic Contiguity} to derive both \textit{wh}-movement and \textit{wh}-in-situ languages; there are two conceivable ways of satisfying the principle. In (9i) below we can see that the \textit{wh}-phrase is in-situ. This means that it is separated from its scope position in C by intervening prosodic IP boundaries. That is, the structure in (8i) violates Prosodic Contiguity. A derivation can change its syntax, via movement, to bring a \textit{wh}-phrase to the periphery, where it can form a short IP together with its CP, as illustrated in (9ii). Alternatively, the derivation’s prosody can be altered by \textit{removing} prosodic boundaries as necessary, in order to derive a
longer IP containing the \( wh \)-phrase, the C head where it takes scope, and so also all material intervening linearly between the two, as illustrated in (9iii).

\[
(9) \quad \begin{align*}
\text{i) } & \text{C & } wh \text{- not in same IP: } (IP \ldots C \ldots) (IP \ldots) (IP \ldots wh- \ldots) \\
\text{ii) Syntactic change: } & (IP \ldots wh-i C \ldots) (IP \ldots) (IP \ldots ti \ldots) \\
\text{iii) Prosodic change: } & (IP \ldots C \ldots \ldots \ldots \ldots wh- \ldots)
\end{align*}
\]

In the remainder of this paper, we present \( wh \)-in-situ languages as satisfying Richards’ principle by this second means. We therefore show how \( wh \)-in-situ languages’ prosody is changed to satisfy prosodic contiguity.

2.3 Japanese \( wh \)-prosody

We now begin analysis of data from \( wh \)-in-situ languages in earnest, and in doing so present data to correspond to the syntactic and prosodic structures involved in a derivation. To illustrate the preceding theoretical discussion in the light of more established research, we briefly present analysis from a volume of research on the prosody of Japanese dialects (Ishihara 2003, see also Smith 2013, 2014), much of which has been foundational to the development of prosodic contiguity and proposals similar to it.

Figure 1 exemplifies the prosody of the Tokyo dialect studied by Ishihara². Each word is parsed into an IP with a noticeable rise and fall in pitch.

\[
(10) \quad \text{Naoya-ga nanika-o nomiya-de non da (Ishihara 2003: 28a)} \\
\quad \text{Naoya-NOM nanika-ACC bar-LOC drank} \\
\quad \text{‘Naoya drank something at the bar’}
\]

² It should be noted that Ishihara also notes how prosody is predicted to disambiguate the scope of an embedded \( wh \)-phrase in Japanese. Simpler examples are presented here for ease of illustration.
In an interrogative sentence, however, this pitch contour is suppressed on words which come between the *wh*-phrase and the question particle (i.e. the C head):

(11) Naoya-ga nani-o nomiya-de nonda no? (Ishihara 2003: 28b)

Naoya-NOM what-ACC bar-LOC drank Q

‘What did Naoya drink at the bar?’

Richards takes this contrast as a foundational point in his development of prosodic contiguity. His account, which we adopt before extending to additional languages, derives the prosodic structure of (11) from that of (10), by removing the boundaries that would create individual IPs between the *wh*-phrase and question particle, thereby satisfying his proposed principle.

(12) **Declarative prosody:** (IP Naoya-ga) (IP nanika-o) (IP nomiya-de) (IP nonda)

**Interrogative prosody:** (IP Naoya-ga) (IP nani-o nomiya-de nonda no)
As mentioned, each IP in Tokyo Japanese is regularly taken to be associated with a noted pitch contour. The process of pitch suppression associated with an interrogative sentence is then proposed to result from running the relevant structure into a single IP, with the lack of any prosodic peaks signifying there is no new, separate IP.

We have now developed an understanding of how prosody changes in monoclausal declarative and interrogative contexts. The following sections consider cases of embedded \textit{wh}-phrases, which raise the issue of prosodic disambiguation of \textit{wh}-scope.

3. **Telugu (Dravidian)**

We now present research which shows that the framework built to account for Japanese prosody extends to other languages. Consider first Telugu, which is a head-final, SOV language (Krishnamurti & Gwynn 1985), like Japanese, though a member different language family.

3.1 **Telugu declarative prosody**

Example (13), together with Figure 3, shows the characteristic prosodic structure of a declarative (ie. non-question) Telugu sentence.

\begin{center}
\begin{tabular}{l}
(13) Kalyani Naraayana gaarelu waṇḍææɖu ani anukondi \\
Kalyani._FEM Naraayana._MASC doughnuts cooked._MASC that think._FEM
\end{tabular}
\end{center}

'Kalyani thinks that Naraayana cooked lentil doughnuts'
Each word has a notable rising prosodic contour, with successive structure characterized by prosodic downstepping, in which this contour repeats at a lower pitch. Thus, the highest pitch in *Naraayana & gaarelu* is lower than that on the previous word, but all DPs have the same contour. There is then a rise on the embedded verb, a pause before the complementiser *ani*, and an utterance-final fall containing the matrix verb. Considering the syntactic structure of this example allows us to propose a mapping between syntax and prosody which derives this prosody.

(14)

We hence propose to analyse Telugu prosody as depending on XPs’ right edges to instantiate the boundaries between intonational phrases. That is, if the right edge
of every XP (ie. CP, TP, VP, and DP) ends an IP, a prosodic constituency is derived in which every word is parsed into its own IP.

\[(15) \quad (IP\ Kalyani)\ (IP\ Naraayana)\ (IP\ gaarelu)\ (IP\ wan\&\&u)\ (IP\ ani)\ (IP\ anukondi)\ (\&C)\]

This accounts for the repeated prosodic contour observed on the first four words: each IP is simply associated with a rising pitch. We suggest the sentence end is independently associated with a fall in pitch, thus accounting for the prosody of the final two words. With a simple account of declarative Telugu prosody established, we begin analysis of the prosody associated with wh-phrases.

### 3.2 Telugu wh-prosody

The sentence in (16) is an example from Telugu of the discussed wh-scope ambiguity inherent to wh-in-situ languages.

\[(16) \quad Kalyani\ Naraayana\ yemi\ wan\&\&u\ ani\ anukondi\]

\[\text{Kalyani}.\text{FEM}\ \text{Naraayana}.\text{MASC}\ \text{what}\ \text{cook}.\text{3MASC}\ \text{that}\ \text{think}.\text{3FEM}\]

Matrix wh-scope: ‘What does Kalyani think Naraayana cooked?’

Embedded wh-scope: ‘Kalyani thinks about what Naraayana cooked’

The tree structure in (17) shows the constituents which are predicted under Richards’ prosodic contiguity to be parsed into a single IP. For prosody to express interpretation of (16) with embedded wh-scope, the embedded wh-phrase, verb, and complementiser should be parsed into one IP, while for matrix wh-scope, we predict the resulting IP to contain additional structure, including the matrix verb and (null) complementiser.
The predicted prosodic constituency is indeed observed. Figures 4 and 5 below show the prosody of the sentence as interpreted with matrix or embedded wh-scope respectively. Consider first how the sentence is pronounced with matrix wh-scope, that is, as a direct question, which is predicted to have the prosodic constituency in (18).

(18) (IP Kalyani) (IP Naraayana) (IP yemi wandææɖu ani anukondi øC)

Figure 4. Telugu matrix wh-scope prosody
Noticeably, the embedded verb no longer has the rise in pitch that it had in the declarative: its prosody is flattened and, furthermore, this flattening continues to the end of the sentence, applying over the complementiser and matrix verb, thus causing the loss of the utterance-final fall. We analyse this, in the established manner, as the result of erasing the prosodic boundaries between the \(wh\)-phrase and the matrix C, the result being the creation of one long IP, as in (18), in which, following the rising tone on the first word, a high pitch is maintained.

When considering the prosody of the sentence with embedded \(wh\)-scope interpretation, though, notice that this domain of flattening does not continue to the sentence’s end: the final prosodic fall is maintained.

![Figure 5. Telugu embedded \(wh\)-scope prosody](image)

This is accounted for under exactly the same principles as derived the prosody associated with matrix \(wh\)-scope. Boundaries between IPs are again erased, but in this case only between the \(wh\)-phrase and the embedded C. The result is an IP which has the same higher, flatter pitch, but is somewhat shorter.

\[
\text{(19) } \text{(IP Kalyani) (IP Naraayana) (IP yemi \text{ wandæeðu} ani) (IP anukondi) (Ø C )}
\]

This allows for the remaining structure to the right to show the sentence-final fall - the expected prosody, as observed earlier in the derivation of the declarative.
4. Khmer (Cambodian | Mon-Khmer)

The study now turns to Khmer, which we present as a counterpart to Telugu, as a head-initial, SVO language which can parse larger syntactic structures into a single, long IP. In the following, we show that Khmer *wh*-prosody cannot be derived in a parallel manner to what is observed for Telugu or Japanese above. We identify an inherent issue for prosodic contiguity and propose that the need to disambiguate *wh*-scope accounts for the novel prosody observed in Khmer. As with Telugu, analysis of Khmer grammar is our own, though compatible with existing studies of the language’s syntax and phonology (Haiman 2011, Butler 2014) which the reader may consult for broader, more detailed analysis.

4.1 Khmer Declarative prosody

The following is an example of a declarative Khmer sentence analogous to the cases presented above. With Khmer being an SVO language, the embedded object is found at the very end of the sentence. A corresponding tree structure is given in (21).

(20) Boran duŋ-tʰaa Rany rian barañ
Boran knows Rany learns French
‘Boran knows Rany is learning French’

(21) 

\[
\begin{array}{c}
\text{CP} \\
\text{TP} \\
\phi \\
\text{DP} \quad \text{T} \quad \text{VP}
\end{array}
\]

\[
\begin{array}{c}
\text{Boran} \\
\text{V} \\
\text{C} \\
\text{TP}
\end{array}
\]

\[
\begin{array}{c}
\text{duŋ-tʰaa} \\
\phi \\
\text{DP} \quad \text{T} \quad \text{VP}
\end{array}
\]

\[
\begin{array}{c}
\text{Rany} \\
\text{V} \\
\text{DP}
\end{array}
\]

\[
\begin{array}{c}
\text{rian} \\
\text{barañ}
\end{array}
\]
The sentence’s prosody is given in Figure 6. A level pitch is found throughout, though with a gradual fall towards the end of the utterance.\(^3\)

![Figure 6. Khmer declarative prosody](image)

We take this as reason to propose the following mapping between the sentence’s syntactic and prosodic constituency:

\[
(22) \quad \begin{array}{c}
\text{[CP Boran } \text{duŋ-tʰaa } \text{[CP Rany riən } \text{baraŋ]} \\
\text{[IP Boran } \text{duŋ-tʰaa } \text{Rany riən } \text{baraŋ]}
\end{array}
\]

The entire sentence, containing both matrix and embedded clauses, is parsed into one IP. This constituency can be derived by analysing the language’s XPs, except for DPs, as using their right edge to mark prosodic boundaries. As the language is SVO, and in this case involves no movement or more complex structure, the right edge of every relevant phrase corresponds to the right edge of the utterance. This configuration raises a potential issue for prosodic contiguity: are there consequences if it is always satisfied?

\(^3\) The speaker of the Khmer examples is female (in contrast to the speaker of the Telugu examples who is male). Spikes in the pitch track are caused by consonants: [b] and [d] are ingressive, while [ɾ] may be a trill.
4.2 Khmer wh-prosody

Under Richards’ principle of prosodic contiguity, Khmer is correctly predicted to not move wh-phrases. No matter its position, any wh-phrase in a derivation like (22) will come to be contained in the same IP as its corresponding C head because the entire sentence forms only one IP. Above, it was noted that prosodic contiguity derives different prosodic structures in Telugu which serve to disambiguate the scope of an embedded wh-phrase. For Khmer no such correlation can be predicted, but, as we now show, disambiguation continues to rely on prosody. (23) is an equivalent sentence to (22), save the presence of an embedded wh-object, thus giving rise to two possible interpretations:

(23) Boran duŋ-tʰaa Rany rian ai

Boran knows Rany learns what

Matrix wh-scope: ‘What does Boran know Rany is learning?’
Embedded wh-scope: ‘Boran knows what Rany is learning’

We present first the the prosody of (23) as interpreted with embedded wh-scope, noting the striking similarity to the prosody of the declarative. Figure 7 shows the same fairly level prosody with a gradual fall at the sentence end:

Figure 7. Khmer embedded wh-scope prosody
However, the prosody of the sentence if interpreted with matrix *wh*-scope is distinct. In this case, shown in Figure 8, prosody is again relatively level, though at a higher pitch, and with a distinct rising tone on the *wh*-phrase:

![Figure 8. Khmer matrix *wh*-scope prosody](image)

We account for this novel prosody by proposing that in this context Khmer makes use of an independent intonational morpheme which we also observe on the language’s polar questions. An analysis parallel to this is found for French, an optional *wh*-movement language, where *wh*-in-situ and polar questions share a particular intonation which is not found on questions derived by *wh*-movement (Cheng and Rooryck 2000, Oiry and Demirdache 2006). This morpheme obviates the need for further prosodic disambiguation. With disambiguation of the two interpretations of (23) made possible by the tone associated with this morpheme, there is no need for *wh*-prosody to differ from declarative prosody, though, indeed, the prosodic constituency derived in (22) would not provide a means of doing this in the first place.

Thus, a new prosodic pattern related to *wh*-scope is identified, one which prosodic contiguity and related theories do not predict. The pattern arises, we suggest, because of a global principle which holds over the derivation’s output, ensuring it can be disambiguated from other, related derivations.
5. Conclusion

The existence of both \(wh\)-movement and \(wh\)-in-situ cross-linguistically raises significant questions about grammar and language design. Even if it is assumed that in-situ \(wh\)-phrases undergo covert movement to derive a logical form akin to moved \(wh\)-phrases, it remains to be explained what allows in-situ \(wh\)-phrases to occur in overt syntax. The prosody of languages like Japanese and Telugu allows prosodic contiguity to indicate the position where \(wh\)-phrase takes scope. Prosodic contiguity, therefore, is one means of allowing \(wh\)-phrases to stay in-situ. However, in Khmer we do not see the intonation associated with prosodic contiguity, even though the language is \(wh\)-in-situ. In Khmer the mapping between syntax and prosody always satisfies prosodic contiguity and this means the principle cannot distinguish between embedded and matrix \(wh\)-scope. In accounting for the prosody observed in the language, we argued that disambiguation is achieved through the use of an intonational morpheme associated with polar questions. This morpheme keeps derivations with differing semantic interpretation from being realised with identical syntax and phonology, and so creates a second context in which \(wh\)-phrases do not need to move.

We have seen that principles governing the syntax-prosody interface can successfully derive the contrasting syntax of \(wh\)-movement and \(wh\)-in-situ languages. However, there is more than one kind of prosodic realisation which can disambiguate \(wh\)-scope, and therefore Richards’ condition on \(wh\)-prosody undergenerates. The question of why \(wh\)-in-situ languages adhere to more than one set of prosodic principles in disambiguation of \(wh\)-scope is an open question for further research, but we close by noting that these languages remain alike in relying on prosody, rather than movement, to disambiguate \(wh\)-scope.

Acknowledgements

We would like to thank Adam Albright, Sudheer Kolachina, Norvin Richards, and Jen Smith for valuable comments and discussion as this project has developed. Further thanks are owed to the speakers of Khmer and Telugu we have worked
with, Rany Bradley, Sudheer Kolachina, Pandu RV Kuchibhotla, and Padma Parakala.

References
Haiman J 2011 Cambodian: Khmer Amsterdam/Philadelphia: John Benjamins
Huang C-T J 1982 ‘Move WH in a language without WH movement’ The Linguistic Review 1 369-416
Ishihara S 2003 Intonation and Interface Conditions PhD thesis, MIT
Krishnamurti BH and Gwynn JPL 1985 A Grammar of Modern Telugu OUP: Delhi
Oiry M and Dermidache H 2006 ‘Evidence from L1 Acquisition for the Syntax of Wh-Scope Marking in French’, in Torrens V and Escobar L (eds) The Acquisition of Syntax in Romance Languages John Benjamins: Amsterdam
Reinhart T 1998 ‘Wh-in-situ’ in the framework of the Minimalist Program’ Natural Language Semantics 6 29-56
Richards N 2011 Uttering Theory Manuscript, MIT
Smith J 2013 ‘Fukuoka Japanese wh prosody in production and perception’ Lingua 124 96-130