Inner Speech in Working Memory During Silent Reading:

Effects of Articulatory Suppression on Anticipated Lexical Stress

Irene Tang

Swarthmore College

Faculty Advisor: Daniel Grodner

May 2019

Submitted to the faculty of the Departments of Psychology and Linguistics at Swarthmore College in partial fulfillment of the requirements for the degree of Bachelor of Arts.

Table of Contents

Abstract

As we silently read words on a page, we use an "inner speech" that includes rhythm and inflections as if we are reading aloud. We also use an inner speech when performing non-reading tasks such as remembering verbal materials, problem solving, mental calculation, and decision making. Present research has not yet determined whether these two uses of inner speech occupy the same cognitive resources. This paper presents findings from an eye-tracking study designed to explore whether the inner speech of silent reading occupies the same resources in working memory as other mental tasks. In the first part of the experiment, participants read stress-alternating homographs (e.g., PREsent, preSENT) embedded in limericks, which compelled them to initially expect the incorrect prosody of the homograph and thus encounter a reading cost. As they read, participants also performed articulatory suppression by repeating the word *this* aloud. The goal of this was to use the outer voice to occupy the relevant resources in working memory, thus rendering them unavailable to be used during silent reading—this allows us to see if those resources pertain to the inner voice of silent reading. Participants also completed the 18-item Varieties of Inner Speech Questionnaire (VISQ; McCarthy-Jones & Fernyhough, 2011), a self-assessment survey that assesses one's relationship with inner speech. The goal of this was to see if participants who reported experiencing higher levels of inner speech in their everyday lives would be more susceptible to the reading costs prompted by articulatory suppression during rhythm-mismatching limericks. If the differences in reading costs between reading rhythm-matching and -mismatching limericks vanish while repeating *this*, and if participants who experience higher levels of inner speech are more affected by articulatory suppression, then this would show that these two inner voices occupy the same working memory resources. The study found some preliminary evidence that articulatory suppression can diminish the effect of stress clash on silent reading, as well as some preliminary indication that inner speech occupies working memory and that this machinery overlaps with the inner voice of the rehearsal component of the phonological loop.

Key words: inner speech, silent reading, working memory, phonological loop, articulatory suppression, eye movements

1   Introduction

How inner speech and silent reading intersect has been a topic of interest in cognitive psychology for over a century. At least since Huey's (1908/1968) assertion that "it is perfectly certain that the inner hearing or pronunciation, or both, of what is read, is a constituent part of reading by far the most of people," researchers have tried to discern the nature of the inner voice experienced during silent reading. It is well-established that some form of inner speech is experienced during many cognitive tasks like counting and decision-making due to their use of the phonological loop in working memory, but it is still unknown how or whether the inner speech experienced during such cognitive tasks is related to the inner speech that is experienced during silent reading.

After providing a brief overview of the inner speech phenomenon in Section 1.1, this paper explores the development of research regarding the nature of the inner speech, the cognitive processes involved in silent reading, and whether silent reading occupies the same working memory resources as other mental tasks that use inner speech. Section 1.2 explores the relationship between inner and outer speech. Section 1.3 overviews Baddeley and Hitch's (1974) model of working memory and the phonological loop. Section 1.4 assesses the role of the inner voice in a number of cognitive tasks. Section 1.5 evaluates the evidence and the role of the inner voice in silent reading; specifically, in the cognitive processes of visual word recognition and syntactic parsing. Sections 2-4 respectively present the methods, results, and discussion of the present eye-tracking experiment investigating the cognitive resources involved with silent reading.

1.1 Overview of inner speech

People often notice a non-verbal, internal speech in their heads during cognitive tasks such as silent reading, problem solving, mental calculation, task switching, and decision making. This *inner speech* (sometimes also referred to as *inner voice, inner ear, implicit prosody, subvocalization*, or *phonological recoding*)[1] can be defined as an internal manifestation of vocal language, and it largely resembles the sounds and prosody of the words that are being experienced. It can mimic the voice of the supposed author or character who is speaking (Kosslyn & Matt, 1977), or it can identify as variations of the readers' own voice (Filik & Barber, 2011; Vilhauer, 2016). Inner speech becomes loudest when processing difficult text, such as in beginning readers or while reading advanced material (Coltheart, Besner, Jonasson, & Davelaar, 1979). People can produce inner speech at rates of around ten times that of outer speech (Korba, 1990; see also Anderson, 1982; Foss & Hakes, 1978). Inner speech is quite pervasive—it is estimated to accompany, on average, at least a quarter of our conscious lives (Heavey & Hurlburt, 2008; cf. Klinger and Cox, 1987-88).

1.2 Relationship between inner and outer speech

It was once proposed that inner speech resembles outer speech up until the point where sound exits the mouth—that inner speech is a muffled form of outer speech without the motor processes (Watson, 1913). This is evidenced by the fact that the same throat and mouth muscle movements used during outer speech become active when inner speech occurs, except involuntarily and with minimized intensity such that they are undetectable without the aid of electromyographic recording machines (Sokolov, 1972). Inner and outer speech also appear to

---

[1] This paper does not use this set of terms interchangeably, although in literature they often appear in the same contexts.

share a common cerebral network, both inducing activations in the essential language areas of the brain: Broca's area, Wernicke's area, and the parietal lobe (e.g. McGuire et al., 1996). Indeed, many parallels can be drawn between inner and outer speech.

However, this view of inner speech as merely an absence of audible representation was soon contested by a proposal that inner speech is a function in and of itself that undergoes a profound refining during its process of internalization (Vygotsky, 1934/1978). Support for this view is demonstrated by typical occurrences of inner speech in spite of atypical abilities to perform outer speech. For one, lack of muscle movement itself (e.g. due to paralysis) does not prevent the ability to produce inner speech (Smith, Brown, Toman, & Goodman, 1947), showing that inner speech can occur without articulation. Additionally, Netsell and Bakker (2017) also found that persons who stutter reported being 100% fluent with inner speech and were significantly slower with overt sentence production than with covert production, suggesting that inner speech consumes more abstract mental representations. Thus, it appears that while inner and outer speech share many similarities, the two are distinguished by more than just motor processes.

1.3 Inner speech in working memory

A large body of research is concerned with the role of inner speech in memory and non-linguistic cognition. This work builds upon the multicomponent model of working memory developed by Baddeley and colleagues (Baddeley & Hitch, 1974; Baddeley, 2001). On this model, working memory can be broken down into four sub-components: the *central executive* that manages attention and serves as a coordinator for the other three sub-components, the *visuospatial sketchpad* that deals with visual images, the *episodic buffer* that deals with creating

vivid impressions of emotion-filled memories, and most relevant for present purposes, the

phonological loop that deals with language and auditory information (see Figure 1).

The phonological loop is made up of two sub-components of its own (Baddeley, 1992): a

passive rapidly-decaying *storage buffer* (sometimes referred to as the *phonological store*) which

retains traces of speech in phonological form for one to two seconds at a time, and an active

rehearsal loop (sometimes referred to as the *maintenance loop* or the *articulatory control*) which

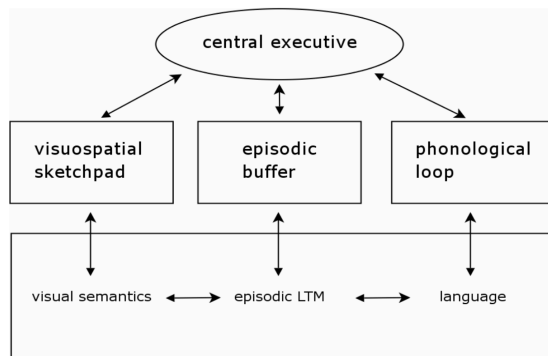refreshes the information in the buffer through silent articulation (see Figure 2).

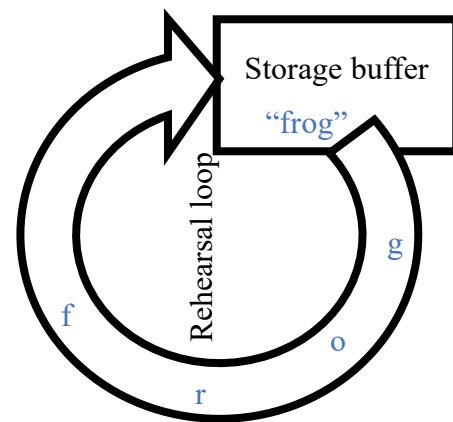
Figure 1: Working Memory, from Baddeley 2003

Figure 2: Phonological Loop

The storage buffer serves as a kind of "inner ear" that hears the "inner voice" produced

by the rehearsal loop. These subcomponents work together but can be independently

manipulated. For instance, Rayner and Pollatsek (1989) demonstrated that sometimes inner

speech can be heard even while articulatory suppression is performed, suggesting that it is

possible for traces of relevant speech to use the storage buffer even when the rehearsal loop is

handicapped with irrelevant auditory information. In addition, the inner ear and voice are with

associated with different areas of the brain: the storage buffer creates activation in Brodmann's

area 44, while the rehearsal loop creates activation in the adjacent Broca's area (Baddeley,

2003).

1.4 Role of inner speech and the phonological loop in various cognitive tasks

The dual-task paradigm is the most frequently-used approach to testing which cognitive

tasks benefit from the phonological loop. In this paradigm, participants perform a task while at

the same time performing articulatory suppression—that is, repeating the same irrelevant syllable

aloud over and over to occupy the rehearsal loop. A great deal of evidence indicates using the

external voice blocks inner rehearsal, thus rendering the inner voice unavailable to be used by the

task. Deterioration of performance on a task when articulatory suppression is performed would

suggest that the task normally benefits from the rehearsal loop, and sheds light into what types of

cognitive work cannot be accomplished without inner speech.

Research has demonstrated that the phonological loop is involved in a number of

cognitive tasks. There is evidence that disruption to the phonological loop causes disruptions to

mental addition and multiplication through disallowing storing partial solutions (Seitz &

Hengsteler, 2002), to task switching by impairing the ability to smoothly offset the "switch cost"

between different tasks (Emerson & Miyake, 2003), to logical or propositional reasoning by

impairing the maintenance for information about logical premises, and to reasoning about false

beliefs in other agents (Newton & de Villiers, 2007), among other tasks.

1.5 Evidence and role of the inner voice in silent reading

A largely independent body of research investigates the role of an inner voice during silent reading. Reading requires (a) visual word recognition and (b) syntactic parsing. There is evidence that each of these processes involves an inner voice.

Visual word recognition (sometimes referred to as *lexical access*) involves identifying and retrieving from memory the semantic meanings of words visually presented on a page. Word recognition invokes two routes to identifying the lexical item (Coltheart, Curtis, Atkins, & Haller, 1993). There is a *direct*, or *lexical, route* accesses meaning from only the visual orthography, as if through a dictionary lookup. There is also an *indirect*, or *phonological*, *route* accesses meaning from its *phonological recoding*, which is a recoding of the orthography into a phonological sound-based mental representation. Whether this recoding during silent reading invokes the same cognitive machinery as the phonological loop is an open question. Regardless, many studies confirm the essential role that the phonological route plays in word recognition.

One family of evidence for phonological recoding during word recognition comes in the form of display change studies. Such studies require the participant to read a sentence in which the target word is initially replaced either by another 'preview' word or by a string of random letters. When the participant reaches that area of the text, the preview word is discreetly replaced by the correct word. This replacement occurs subtly during a saccade, when the eye is functionally nearly blind, so that participants do not notice the change. Pollatsek Lesch, Morris, and Rayner (1992) found that previewing a homophone of target word in the parafovea resulted in faster processing of the target word during the next fixation, compared to previewing a visually-similar but non-homophonic word. This indicates that the phonological representation of the previewed word was accessed and affected the processing of the correct target.

Another experimental technique that has been used to demonstrate the involvement of phonology in word recognition is the fast-priming technique. Like in display change studies, in fast-priming studies the target word is replaced with a non-word until the participant reaches the area of text—except unlike display change studies, a 'prime' word is briefly flashed in the place of the target word. Sereno and Rayner (1992) found that briefly flashing a homophonic prime (for only 36ms, too brief to consciously notice) before switching from the preview word to the actual target word caused readers to identify the target more quickly than if they had been primed with a non-homophone. Rayner, Sereno, Lesch, and Pollatsek (1995) found that such significant reductions in gaze durations extended to being flashed for 36ms with pseudo-homophonic prime words (e.g., *bead* is considered a pseudo-homophone for *bed*) as well. Taken together, these pieces of evidence suggest that creating a phonological recoding of a word helps with word recognition.

Above the level of individual words, there are indications that syntactic parsing is affected by an inner voice. Fodor (1998) proposed that during online sentence comprehension, there is a *prosodic parser* and a *syntactic parser* operating in parallel to identify the syntactic structure of the sentence. The prosodic parser follows the *same-size sister principle* (Fodor, 1998) to break up the sentence into phrases that are roughly the same size, which by default is the most natural prosodic contour ceteris paribus (Fodor, 2002). The syntactic parser then only groups constituents takes these phrase divisions into account when assigning phrase structure to the input. Indirect evidence from attachment ambiguity resolution suggests that perceivers take into account this principle when assigning syntactic structures (e.g., Kitagawa & Fodor, 2006; Swets, Desmet, Hambrick, & Ferreira, 2007).

Meanwhile, Bader (1998)'s Prosodic Constraint on Reanalysis hypothesis argues for phonological recoding's role in recovering from misanalysis of syntactically ambiguous input. This hypothesis claims that it is harder to revise a syntactic misanalysis if it requires the reader to revise the prosodic analysis as well. Support for this claim comes from observing that syntactic reanalysis that requires altering the prosody of an initially built structure is often more difficult than syntactic reanalysis that preserves the initial prosody. For instance, Breen and Clifton (2011) had participants read temporarily ambiguous sentences, such as in (1)[2]:

(1a)   The brilliant rePORT/ABstract was accepted at the prestigious conference.

(1b)   The brilliant rePORT/abSTRACT the best ideas from the things they read.

Brilliant is heavily favored to be an adjective, as in (1a), but can also be a noun, as in (1b). When it is an adjective, the following word must be a noun, but when it is a noun, the following word must be a verb. Participants had more difficulty with continuations like (1b) when reanalyzing from an adjective-noun to a noun-verb, but it was especially pronounced for stress shifting stress-alternating noun-verb homographs (e.g. *abstract*) which required re-prosodifying the input.

While the work reviewed here is consistent with prosodic effects on within-sentence processing, the evidence is somewhat indirect. In order to alter the prosodic properties of the input, the lexical material must be changed as well. This necessarily affects non-prosodic properties as well. For instance, Breen and Clifton's (2011) result could stem from the possibility that the verb forms of non-stress-shifting homographs (like *report*) are relatively frequent compared to the verb forms of stress-shifting homographs.

---

[2] Capitalization indicates stress but was not present in the stimuli.

Indeed, other work suggests that implicit prosody does not affect the processing of individual sentences. Rayner and Pollatsek (1989) argue that phonological coding aids the prosodic structure by strengthening a reader's memory for the text across sentences (see also Gathercole & Baddeley, 1993). "Because short-term memory representations are thought to be predominantly acoustic in nature, the generation of a phonological code strengthens a reader's memory for individual words, giving him time to integrate them into the larger syntactic and semantic context." Slowiaczek and Clifton (1980) provide evidence for this proposal, demonstrating that suppressing the inner voice through articulatory suppression caused impairments to memory for multi-sentence (although not for single-sentence) propositions and inferences. The fact that articulatory suppression did not affect within-sentence interpretation raises the possibility that the phonological loop is independent of the inner voice observed in silent reading of individual sentences. One possibility is that the phonological loop is necessary only for post-interpretive switching processes (Caplan & Waters, 2002; Emerson & Miyake, 2003). Another is that the phonological loop is primarily used to store prosodic cues to discourse and coherence information (e.g. focus and speech act identifiers) that operate across sentences (Rayner & Pollatsek, 1989).

Though evidence exists for the inner voice during reading, it is not clear whether it plays an active role in processing sentences during silent reading or whether it is, to borrow the words of Breen (2014), "simply epiphenomenal—a by-product of the fact that language has been spoken far longer than it has been written" (see Gelb, 1952).

2   Experiment

Some of the strongest evidence that implicit prosody directly affects on-line reading behavior comes from Breen and Clifton (2011). In their first experiment, they found that the metrical structure during silent reading could be disrupted. Participants' eye-movements were tracked as they read limericks where the second line ended with a stress-alternating noun-verb homograph (e.g. PREsent, preSENT). In some of the limericks, the stress-pattern of the correct interpretation of the homograph clashed with the prosodic context set up by the first line of the limerick. It was discovered that readers experienced difficulty upon entering this critical region if they encountered stress patterns that clashed with the expected reading, as demonstrated by measures of their eye movements: first pass time, go-past time, probability of fixating a region, and probability of regressing out of a region give that it was fixated the first time around. Reading costs in clashing contexts were especially noticeable with regards to go-past reading times.

The present study attempts to investigate whether it is specifically prosody that caused this clash effect by targeting the inner voice directly through articulatory suppression. One hypothesis predicts that the inner speech experienced during silent reading is produced by the phonological rehearsal loop in working memory. On this view, the inner speech experienced during silent reading should be the same phenomenon as the inner speech experienced during performance of other mental tasks (e.g. mental calculation), which also rely on the phonological rehearsal loop. Thus, like other such mental tasks, inner speech would be disrupted by articulatory suppression, and there should be no reading differences between words that match

and clash with the lexical stress pattern on the limerick while performing articulatory suppression.

An alternative hypothesis predicts that the inner speech experienced during silent reading is instigated through other means not related to the phonological loop at all. On this view, the inner speech experienced during silent reading should be a separate phenomenon from the inner speech experienced during performance of mental tasks that rely on the phonological rehearsal loop. Thus, inner speech would not be disrupted by articulatory suppression, and there should be apparent reading differences between correctly and incorrectly anticipating a word's lexical stress while performing articulatory suppression.

To evaluate these two hypotheses, we aimed to replicate Breen & Clifton (2011) study with some modifications. Participants read limericks that lead them to correctly or incorrectly expect the metrical structures of stress-alternating homographs that were embedded in the limericks. Following Breen and Clifton's (2011) discovery that contexts that anticipate a weak-strong stress pattern do not elicit different reading effects for stress-alternating homographs, the present study used only limericks with critical regions embedded in strong-weak contexts. In addition, to test the effects of articulatory suppression, the present study incorporated an additional dual-task paradigm such that participants would either tap their foot or repeat "this" aloud as they read.

Method

2.1 Participants

57 undergraduate students at Swarthmore College (42 female, 15 male) ages 17-22 participated in this study (mean age 19.7, SD 1.12). Of these, 10 were eliminated from analysis: eight were eliminated for unrepairable eye-tracking difficulties, and two were eliminated

(according to a preregistered exclusion standard) for answering the comprehension questions with accuracy less than 2.5 standard deviations from the median. Thus, 47 participants (35 female, 12 male) were included in the analysis. All participants were native speakers of English (learned before the age of four and spoken consistently since), and all had normal or corrected-to-normal vision.

This study was approved by the Swarthmore College institutional review board, and all participants provided written informed consent. They were assured that confidentiality and anonymity would be secured and that they could end participation at any time without penalty. At the conclusion of the experiment, they had the option to be fully debriefed. Participants received candy and either psychology course credit or monetary compensation of $10.

2.2 Materials and Design

This experiment followed a 2 (rhythm[3]: match or clash) by 2 (secondary task: foot-tap or articulatory suppression) Latin Square design. Both variables were manipulated within-subjects in a dual-task paradigm. Participants were presented with one version of each of 40 target limericks[4] and also saw 60 filler limericks, following those used by Breen and Clifton (2011).

---

[3] *Rhythm* and *meter* are used interchangeably in this paper, although in poetry they may be distinguished.
[4] Limericks are five-line poetic devices which follow a particular rhyme and meter pattern. A filler limerick is exemplified in (1) (Breen & Clifton, 2011). Lines one, two, and five (e.g. *gray*, *away*, *spray*) obey the same rhyme scheme, as do lines three and four (e.g. *pastor*, *disaster*). Lines one, two, and five each contain five metrical feet (metrical feet are a unit of poetry typically containing one to three syllables), while lines three and four each contain two.

> (1)   While painting the church steeple gray
>   The wind blew our brushes away
>   We said to the pastor
>   "We've had a disaster"
>   He calmly replied, "Let us spray"

The target limericks were divided into six regions, as illustrated by the subscripts in (2). The critical region of the target limericks took place in Region 3, the final word of the second line.

(2)   There once was a penniless peasant, $_1|$

Who couldn't afford a nice $_2|$ present. $_3|$

For so little yield, $_4|$

He worked in the field, $_5|$

And work there was really unpleasant. $_6|$

The critical regions of the target limericks contained one of forty stress-alternating noun-verb or noun-adjective homographs (e.g. PREsent as in 'a gift' and preSENT as in 'to show') selected by Breen and Clifton (2011) from those used by Pitt and Samuel (1990) and the Celex database (Baayen, Pipenbrock, & van Rijn, 1993). For each homograph, the rhythm of the first line set up the expected rhythm of the second line such that the strong-weak version of the homograph (e.g. PREsent) should be expected in the critical region. Half of the critical regions contained a strong-weak homograph (e.g. PREsent) for the rhythm-matching condition, and half a weak-strong homograph (e.g. preSENT) for the rhythm-clashing condition.

For each of the 40 critical homographs, there were two limericks: one used the strong-weak version of the homograph to create a rhythm-matching trial and one the weak-strong version to create a rhythm-clashing trial (see Table 1). The remaining three lines were not of particular experimental interest. Thus, there were 80 target limericks in total, but each participant only saw 40 of the target limericks. No participant saw both versions of a target limerick—that is, no participant saw the same homograph twice.

Table 1: Metrical structure for the first two lines of limericks, two versions presented.

anticipated stress	(w	s)	(w	w	s)	(w	w	s)	(w)
A rhythm-	There	ONCE	was	a	PENN	i	less	PEA	sant
matching	Who	COULD	n't	af	FORD	a	nice	**PRE**	**sent**
B rhythm-	There	ONCE	was	a	PENN	i	less	PEA	sant
clashing	Who	WENT	to	his	MAS	ter	to	**pre**	**SENT**

Participants were also asked to complete one of two secondary tasks while reading the limericks. Half of the trials asked them to repeat the word "this" aloud as they read for the articulatory suppression condition, and the other half to tap their foot as a control for the foot-tap condition.

Following each limerick, participants were asked to judge whether the limerick was 'dirty' or 'clean' in order to ensure that they were reading the limericks for comprehension. None of the limericks contained explicit obscenities, so readers would be required to semantically process the limerick in order to determine whether it was clean or dirty. A 'dirty' limerick was defined for the purposes of this study as containing mild nudity or sexual references, and this definition was relayed to the participant in the instructions screen. Twenty filler limericks were chosen to be 'dirty.' All filler limericks were scraped from the internet by Breen and Clifton (2011).

The 80 target limericks were distributed across two sets of four lists per Latin-Square design. The 60 fillers (40 clean, 20 dirty) were the same on each of the eight lists. Thus, each participant saw 100 experimental trials. Each participant was randomly assigned to one of the eight lists.

After the eye-tracking portion of the study, participants electronically completed a post-experimental survey and the 18-item Varieties of Inner Speech Questionnaire (VISQ; McCarthy-Jones & Fernyhough, 2011). A copy of these can be found in Appendix B. The VISQ was designed to assess one's relationship with inner speech along the dimensions of dialogicality, condensed/expanded quality, evaluative/motivational nature, and the extent to which inner speech incorporates other people's voices. Participants rated each of the 18 statements on a 6-point Likert scale ranging from *Does not apply to me* to *Certainly applies to me*.

2.3 Procedure

All participants were tested individually. Upon providing written consent and partaking in a dominant-eye test, participants were taken into the testing room and asked to have a seat in front of the eye-tracker. The eye-tracker was adjusted to focus on the participant's dominant eye. Participants were then informed that limericks are a type of poetry that are written to be read with a metrical rhythm and rhyme, instructed to read the limericks for comprehension at a normal and comfortable pace, asked to refrain from blinking while reading the limericks, and asked to refrain from moving until the designated break points. They were also informed that each limerick would be accompanied by a secondary task and that they would be asked to judge each limerick as dirty or clean.

After the eye-tracker was well-calibrated (classified as having a "good" track, the highest possible classification), participants read instructions on the screen detailing the procedure and the definition of "dirty" for the purposes of this study. Following the instructions were four practice trials that participants completed with onlooking from the experimenter. Two of these were dirty limericks and two clean. For the first two practice trials, there was no secondary task. The next two required participants to engage in the foot tapping and repeating "this" secondary

tasks respectively. Participants were told to engage in the secondary task repeatedly without stopping. Feedback about the correct judgement of whether the limerick was dirty or clean was shown on the screen after each practice limerick, along with its reasoning.

The experiment then began. Prior to each trial, the monitor presented instructions for the secondary task, displaying either "please begin tapping your foot" or "please begin repeating 'this.'" Once they were done reading this instruction, participants pressed any button on a game controller to continue to a blank screen containing only a blue dot where the first letter of the limerick would appear. Limericks were initiated on the screen once the participant fixated their gaze on the blue dot. After reading the limerick, participants pressed any button on the game controller to bring them to a screen that displayed instructions to stop performing the secondary task. Following the pressing of another button, participants saw a screen that displayed the comprehension question "Was this limerick clean or dirty?" and answered by pressing one of two triggers to answer "yes" or "no." No feedback was given about whether participants answered this question correctly or incorrectly.

The eye-tracker was re-calibrated at each of the two break points, which occurred after 33 and 67 experimental trials. The experimenter was present throughout the experiment to make sure participants consistently performed the secondary task and admonished them when they did not. Limericks were presented in 24-point Times New Roman font on a Dell monitor. Eye movements for only the dominant eye were recorded using an EyeLink 1000+ system controlled by a PC running Experiment Builder software. All participants completed the study in one sitting, with most completing within an hour.

At the completion of the eye-tracking experiment, participants were taken into a second room to complete an exit survey and the Varieties of Inner Speech Questionnaire (VISQ;

McCarthy-Jones & Fernyhough, 2011). All participants completed the surveys within a total of fifteen minutes.

3   Results

3.1 Eye-tracking study

The analyses and exclusion criteria for this study were preregistered on AsPredicted.org (ID#20414). Excluding the participants who were eliminated from analysis, the remaining participants' mean and median comprehension question accuracy scores were 88.55% and 89.50%, respectively. The questions were subjective (one participant commented that they would ordinarily consider death to be dirtier than nudity), so some variation in responses is expected; but nonetheless, the high accuracy suggests that participants were engaged in reading the limericks.

Eye-tracking data was cleaned prior to analysis using various software, following Breen and Clifton (2011). Because eye movement data was not perfectly recorded and tracking quality generally declined over the course of an experiment, misaligned vertical fixation locations were corrected and outliers were removed using fix_align.R software developed by the UMass eye-tracking laboratory (Cohen, 2012). Blinks and track losses were cleaned from all trials, and trials with blinks or track losses in the critical region were eliminated altogether using RoboDoc.py (Staub, Clifton, & Harris, 2015). Out of 1600 total collected target trials, 129 were excluded—leaving a total of 1471 trials to be analyzed, representing 92% of the trials. On the assumption that fixations greater than 800ms or less than 80ms do not represent normal acquisition of

information from text (Rayner et al., 1989), such fixations were eliminated from analysis using

EyeDry.c (Clifton, 2013).

Following Breen and Clifton (2011), the limericks were divided into six observable

regions for analysis, as demonstrated in (2):

> (2)   There once was a penniless peasant, 1|
>
> Who couldn't afford a nice 2| present. 3|
>
> For so little yield, 4|
>
> He worked in the field, 5|
>
> And work there was really unpleasant. 6|

The primary experimental interest laid in Region 3, but due to the possibility that reading times

for Region 3 might be affected by the nature of the region leading up to it, we also looked at the

measures in Region 2.

Following Breen and Clifton (2011; see also Rayner, 1998; Rayner et al., 1989), the

following standard eye-tracking measures were designated for analysis in Regions 2 and 3: (1)

first pass time, the sum of all fixation durations made from first entering the region to first

leaving it; (2) go-past time, the sum of all fixation durations made from first entering the region

to first leaving it to the right; (3) probability of fixating the region; and (4) probability of

regressing out of a region given that it was fixated during first pass.

In their original study, Breen and Clifton (2011) found reliable effects of stress clash for

their weak–strong stimuli in these measures in these regions. Their results are reproduced in

Table 2 as follows (only the Weak-Strong condition was used in the present study):

Table 2: Mean eyetracking measures with standard errors in parentheses, from Breen and Clifton (2011)

Metrical Condition	Region 2 Lexical target		Region 3 Lexical target	
	Consistent	Inconsistent	Consistent	Inconsistent
Proportion fixated				
Strong–weak	.98 (.01)	.98 (.01)	.92 (.02	.89 (.02)
Weak–strong	.97 (.01)	.96 (.02)	.90 (.02)	.94 (.01)
Regressions out (proportion)				
Strong–weak	.03 (.01)	.01 (.01)	.28 (.03)	.28 (.03)
Weak–strong	.02 (.01)	.03 (.01)	.24 (.03)	.33 (.03)
First pass (ms)				
Strong–weak	867 (21.0)	815 (21.3)	286 (9.6)	266 (9.4)
Weak–strong	887 (21.8)	956 (23.6)	287 (10.8)	288 (11.1)
Go-past (ms)				
Strong–weak	906 (23.0)	857 (25.5)	445 (21.3)	434 (22.8)
Weak–strong	909 (23.6)	1008 (24.7)	451 (24.4)	535 (31.8)

Their results show that readers were less likely to fixate on (that is, more likely to skip) the critical word in Region 3 when it was consistent with the preceding context than when it was inconsistent, suggesting that it was more likely to be identified while reading the previous word in Region 2 (Rayner, 1998). Likewise, their participants demonstrated a marginally higher probability of regressing out of Region 3 when its reading was inconsistent with the context. Moreover, go-past times in Region 3 significantly increased in inconsistent contexts. However, first pass reading times in Region 3 showed little indication of difficulty in the critical region in inconsistent contexts. It should be noted that these cost effects were seen only when the critical word followed a weak-strong stress pattern and not a strong-weak pattern—perhaps due to non-prosodic factors such the relative acceptability and frequency of English speakers practicing right-ward stress shifts (Grabe & Warren, 1995).

Looking at reading patterns in Region 2 of the strong-weak stimuli, Breen and Clifton (2011) found that first pass and go-past reading times were slightly faster in the inconsistent

condition compared to the consistent condition. At a glance this seems counter-intuitive, but the authors note that these differences are difficult to interpret given the differences in content in this larger region. For the weak-strong stimuli, first pass and go-past reading times were longer in the inconsistent condition, as expected if critical word in Region 3 is being previewed in the parafovea (Rayner, 1998). Readers almost always fixated in Region 2 and rarely regressed out, presumably because this region spanned most of the second line of text.

The present experiment applied these same four standard eye-tracking measures for analysis. After applying the preregistered exclusion criteria, the means of all dependent measures across conditions were similar. However, participants only fixated the critical one-word Region 3 in the present study ~50% of the time in all conditions, compared with ~90% for Breen and Clifton (2011). Participants also appeared to read at unexpectedly faster rates, averaging ~140ms and ~300ms respectively for first pass and go-past times, compared with ~280 and ~450 for Breen and Clifton (2011). It is unclear whether this was due to difficulties with calibrating the eye-tracker, adjusting its initial settings, a misunderstanding regarding Breen and Clifton's exclusion criteria, complications with the data software, issues specific to our participant pool, or something else.

Thus, a series of exploratory analyses restricting the data to just those participants who fixated the critical region on at least 50% of experimental trials. After performing this adjustment, the calculated measures become more comparable with those reported by Breen and Clifton (2011). While it is important to exercise caution in interpreting these analyses because they were not preregistered, they can serve to establish appropriate exclusion criteria for future studies using our methodology and apparatus.

For each of Regions 2 and 3, a 2x2 ANOVA was performed with participants by crossing clash type (clash or match) and the secondary task (foot tap or articulatory suppression). For proportional measures (probability of fixation, probability of regressing out) data was transformed using an empirical logit to satisfy the normality assumptions of ANOVA (Creel & McNeil, 2002).

First pass time

For Region 3, there was no main effect of clash type ($F(1,31)=0.820$, $p=0.372$) nor secondary task ($F(1,31)=0.614$, $p=0.439$). There was also no significant interaction between clash type and secondary task ($F(1,31)=2.063$, $p=0.161$). For Region 2, there was no main effect of clash type ($F(1,31)=0.037$, $p=0.850$). However, there was a significant main effect of secondary task ($F(1,31)=8.380$, $p=0.007$), whereby the critical region was more difficult to process in the rhythm-clash condition. No interaction was found between clash type and secondary task ($F(1,31)=0.449$, $p=0.508$). Figure 3 depicts average first pass reading times for Regions 2 and 3 across the four conditions.

Like Breen and Clifton (2011)'s weak-strong condition, we found no significant differences in first pass reading times between matching and clashing prosodic contexts in either region. However, Breen and Clifton's (2011) first pass measures in Region 2 trended towards demonstrating longer reading times in rhythm-clashing contexts, whereas our measures showed slightly faster reading times in rhythm-clashing contexts. As mentioned earlier, differences in reading times for Region 2 are difficult to interpret due to differences in content in this larger region. Additionally, participants demonstrated difficulty reading in the articulatory suppression condition as expected and no reliable difference between rhythm-matching and rhythm-clashing contexts under this condition as predicted.
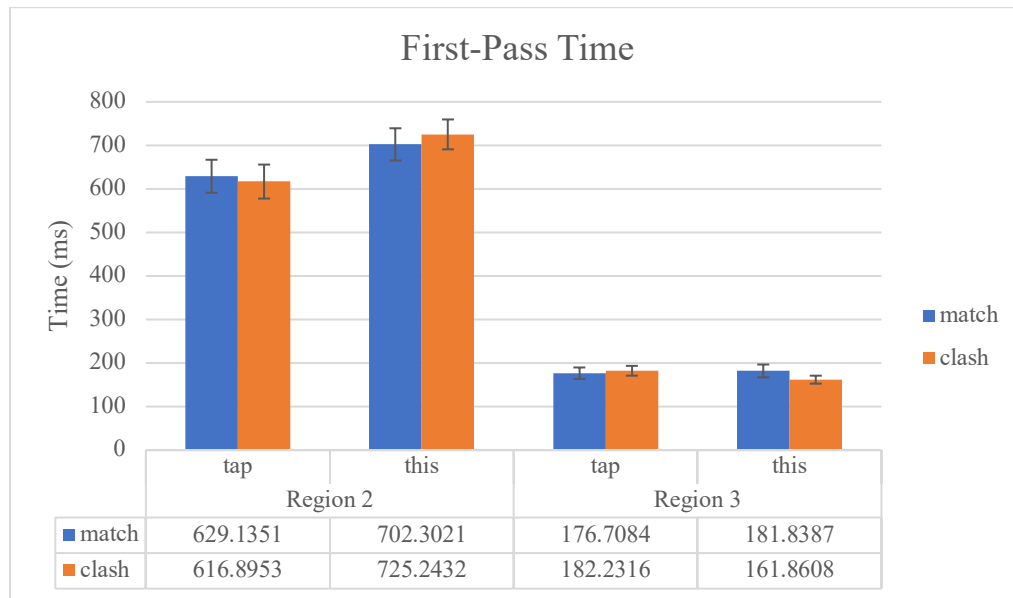
	tap	this	tap	this
	Region 2		Region 3	
■ match	629.1351	702.3021	176.7084	181.8387
■ clash	616.8953	725.2432	182.2316	161.8608

Figure 2: Average first pass reading times

Go-past time

For Region 3, there was no main effect of secondary task ($F_{(1,31)}$=0.021, p=0.885). However, there was a significant main effect of clash type ($F_{(1,31)}$=4.603, p=0.040), whereby the critical region was more difficult to process in the rhythm-clash condition. There was also a marginally significant interaction between clash type and secondary task ($F_{(1,31)}$=3.934, p=0.056), with the effect of clash being smaller in the articulatory suppression condition. For Region 2, there was no main effect of clash type ($F_{(1,31)}$=0.284, p=0.598). However, there was a marginally significant effect found for secondary task ($F_{(1,31)}$=4.040, p=0.053). No interaction was found between clash type and secondary task ($F_{(1,31)}$=0.479, p=0.494). Figure 4 depicts average go-past times for Regions 2 and 3 across the four conditions.

Like Breen and Clifton (2011)'s weak-strong condition, we found significantly longer go-past reading times in clashing prosodic contexts than matching contexts in Region 3, although not in Region 2 where reading times may be affected by non-prosodic factors and parafoveal

intake of information. Additionally, participants demonstrated comparable reading times between rhythm-matching and rhythm-clashing contexts under the articulatory suppression condition. This provides is consistent with the hypothesis that effects of inner speech during silent reading can be reduced through disruption of the phonological loop.
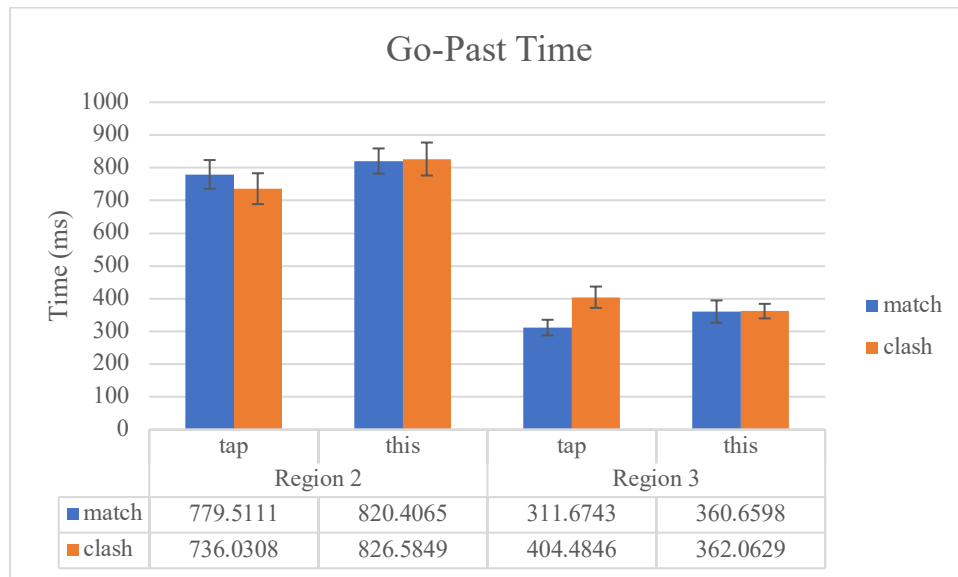

	Region 2		Region 3	
	tap	this	tap	this
■ match	779.5111	820.4065	311.6743	360.6598
■ clash	736.0308	826.5849	404.4846	362.0629

Figure 3: Average go-past reading times

Probability of fixating the region

For Region 3, there was no significant main effect for clash type ($F(1,31)=2.19$, $p=0.149$) nor secondary task ($F(1,31)=0.737$, $p=0.397$). However, there was a significant main effect for the interaction between clash type and secondary task ($F(1,31)=5.2$, $p=0.030$). For Region 2 as well, there was no significant main effect for clash type ($F(1,31)=2.19$, $p=0.149$) nor secondary task ($F(1,31)=0.737$, $p=0.397$). However, there was a significant main effect for the interaction between clash type and secondary task ($F(1,31)=5.2$, $p=0.030$). Figure 5 depicts average probabilities of fixating the region across the four conditions.

After adjusting our exclusion criteria to consider only participants who fixated in the critical Region 3 more than half the time, the probability of fixation in Region 2 became

especially similar to what Breen and Clifton (2011) found. The fact that participants demonstrated comparable reading times for rhythm-matching and rhythm-clashing contexts in both the foot-tap and the articulatory suppression conditions for both regions makes it difficult to assess the effect of articulatory suppression on reading costs, but nonetheless the effects are apparent especially in Region 3.
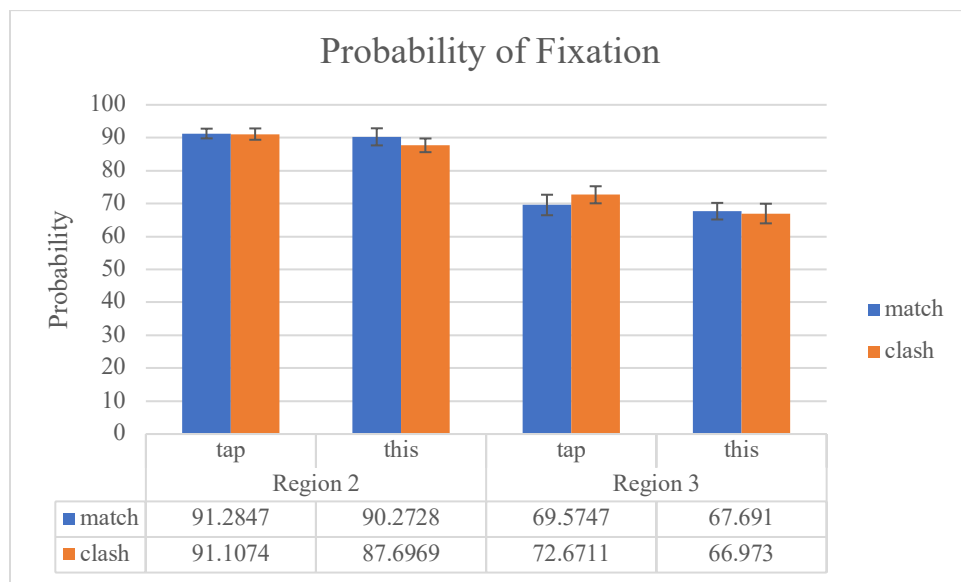

	Region 2		Region 3	
	tap	this	tap	this
■ match	91.2847	90.2728	69.5747	67.691
■ clash	91.1074	87.6969	72.6711	66.973

Figure 4: Average probability of fixation

Probability of regressing out of the region

For Region 3, there was no significant main effect of secondary task ($F(1,31)=1.314$, $p=0.260$). However, there was a trend toward a main effect of clash type ($F(1,31)=2.727$, $p=0.109$), whereby the critical region elicited more regressions in the rhythm-clash condition. No interaction was found between clash type and secondary task ($F(1,31)=2.202$, $p=0.148$). For Region 2, there was no significant main effect of clash type ($F(1,31)=1.559$, $p=0.221$). However, there was a significant effect found for secondary task ($F(1,31)=4.527$, $p=0.041$).). No interaction was found between clash type and secondary task ($F(1,31)=0.052$, $p=0.821$). Figure 6 depicts average probabilities of regressing out of the region across the four conditions.

Like Breen and Clifton (2011)'s weak-strong condition, we found greater probability of regressing out of the critical Region 3 in clashing prosodic contexts than in matching contexts, although not in Region 2. The probability of regressing out of Region 2 was already small, perhaps due to the fact that this region spans across most of the second line so regressions from a word in this region would still land within the region. Regression from Region 2 back into Region 1 also likely do not have a bearing on assessing the reading costs of the critical word in Region 3. Numerically, participants demonstrated comparable probabilities of regression from Region 3 between rhythm-matching and rhythm-clashing contexts under the articulatory suppression condition, but increased regressions in clashing contexts in the foot tap condition. This interaction was not significant but is potentially consistent with the hypothesis that the inner speech during silent reading can be taken away through disrupting the phonological loop.
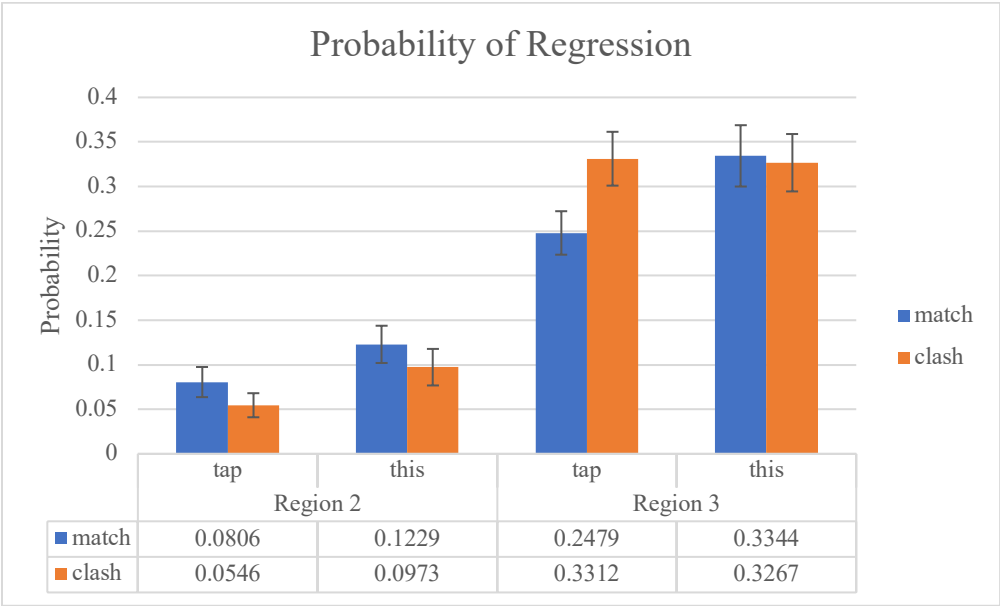

Probability of Regression

	Region 2		Region 3	
	tap	this	tap	this
match	0.0806	0.1229	0.2479	0.3344
clash	0.0546	0.0973	0.3312	0.3267

Figure 5: Average probability of regression

3.2 Varieties of Inner Speech Questionnaire

Following McCarthy-Jones & Fernyhough (2011), the 18 items asked by the Varieties of Inner Speech Questionnaire (VISQ; McCarthy-Jones & Fernyhough, 2011) were classified as relating to one four factors: Condensed Inner Speech (items 1, 7, 8, 14, and 15), Dialogic Inner Speech (2, 6, 10 and 13), Other People in Inner Speech (items 3, 4, 5, 12, and 16) and Evaluative/Motivational Inner Speech (9, 11, 17, and 18) (see Appendix B for a copy of questionnaire items).

Responses to each item were also converted to a 6-point scoring scale, with items 7 and 15 being reverse scored (following McCarthy-Jones & Fernyhough, 2011).

1: Certainly does not apply to me

2: Possibly does not apply to me

3: If anything slightly does not apply to me

4: If anything applies to me slightly

5: Possibly applies to me

6: Certainly applies to me

Participants' average scores for each factor was taken as a measurement of that respective aspect of their inner speech. Due to the unclear differences in measures across conditions and time constraints, this data is pending analysis.

4   Discussion

The present study found some preliminary evidence that articulatory suppression can diminish the effect of stress clash on silent reading. These findings replicate the finding that readers experienced difficulty when they read homograph words using a prosody that clashed with expected prosody. The effect was apparent in slightly longer first pass and go-past times, and a slightly greater probability of fixating and regressing out of the critical region and the region leading up to it for some of our conditions.

This study also provided some preliminary indication that inner speech occupies resources in working memory and that this machinery overlaps with the inner voice of the rehearsal component of the phonological loop. Three of our measures (go-past time, probability of fixating a region, and probability of regressing out of a region given that it was fixated during first-pass), suggested that there may be no reading differences between words that match and clash with the lexical stress pattern on the limerick while performing articulatory suppression.

An important issue for future investigation will be to replicate this experiment after resolving methodological issues with calibration of the eye-tracker. Even though the eye-tracker was calibrated to have a "good" track at the start of each experimental block, the track sometimes appeared to shift slightly as the trials progressed. This was demonstrated by an increased difficulty in activating the trigger to display the limerick in the later trials of each block. The fix_align.R software (Cohen, 2005) hopefully alleviated some of these unwanted shifts in calibration, but it cannot be certain that it fully took care of the issue. There may have been additional settings regarding the eye-tracker that need to be adjusted, such as the minimal amount of time needed for a pause in eye movement to be considered a fixation. It is possible

that the present experiment consistently demonstrated faster reading times compared to those in Breen and Clifton's (2011) due to the fact that the criteria for minimum fixation duration were programmed differently. This might even account for the issue where participants in the present study appeared to fixate in the critical Region 3 only approximately half as often as participants did in Breen and Clifton's (2011) study—it is possible that fixations in this region were very brief (and thus unregistered by the eye-tracker) due to having parafoveally processed the word before reaching the area (Rayner, 1992).

Another interesting avenue for future investigation could be the experience of increased arousal during the experiment due to the experimenter being in the room, despite being silent and out of sight in the back. This could have caused enhanced or impaired reading abilities due to evaluation apprehension theory, which proposes that feelings of concern about evaluation occur when in the presence of others (Cottrell, 1972). This increased arousal can increase performance on the task if the participant is very familiar or well-prepared for the task, or it can decrease performance if they are not. The aim of having the experimenter present was to ensure that participants were reading for comprehension and performing the secondary task correctly, but while this ensured correct participation practices, it could have introduced heightened perception of evaluation that could affected the participants' natural abilities to read the limericks. Indeed, it was sometimes necessary for the experimenter to remind the participants to read at their normal pace and to perform the indicated secondary task every time. Future replications of this study should make further considerations regarding how to ensure that participants are performing the task correctly without drawing attention to evaluation.

Beyond methodological improvements, we may aim to use the 35-item Varieties of Inner Speech Questionnaire–Revised (VISQ-R; Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones,

& Fernyhough, 2018) instead of the more dated 18-item Varieties of Inner Speech Questionnaire (VISQ; McCarthy-Jones & Fernyhough, 2011) to assess participants' relationship with inner speech. Both versions of the questionnaire assess respondents' experience with inner speech through questions about inner dialogue, but the VISQ-R version incorporates an additional *positive/regulatory* aspect to the original assessment scale in the VISQ (which previously only included dimensions of dialogicality, condensed/expanded quality, evaluative/motivational nature, and presence of other people). This revised scale has demonstrated exploratory improvements in the previous questionnaire's ability to link qualities of inner speech to pathological concerns, thus suggesting that it can be repurposed for the present experiment to serve as an even more comprehensive assessment of inner speech.

5   Acknowledgements

6   References

Alderson-Day, Ben & Mitrenga, Kaja & Wilkinson, Sam & McCarthy-Jones, Simon &

Fernyhough, Charles. (2018). The varieties of inner speech questionnaire - Revised

(VISQ-R): Replicating and refining links between inner speech and psychopathology.

Consciousness and cognition. 65. 48-58. 10.1016/j.concog.2018.07.001.

Anderson, R. D. (1982). Speech imagery is not always faster than visual imagery. *Memory &

Cognition*, 10, 371-380.

Baddeley, A. D. (1992). Working memory. *Science*, 255, 556-569.

Baddeley, A. D. (2001) Is working memory still working? *American Psychologist*, *56*(11), 851-

864).

Baddeley, A. D. (2003) Working memory: looking back and looking forward. *Nature Reviews

Neuroscience* 4:829-839.

Baddeley, A. D., & Hitch, G. (1974) Working memory. In G. Bower (Ed.), *The psychology of

learning and motivation* (Vol. 8, pp.47-90). New York: Academic Press.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. Fodor

& F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 1-46). Dordrecht: Kluwer.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The *CELEX lexical database* (CD-

ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania

Breen, M., & Clifton, C. Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent

reading. *Journal of Memory and Language, 64*(2), 153-170. doi: DOI

10.1016/j.jml.2010.11.001

Breen, M., & Clifton, C. Jr, (2014). Stress matters revisited: A display change experiment.

      Quarterly Journal of Experimental Psychology. DOI: 10.1080/17470218.2013.766899

Clifton, C. (2013). EyeDry [Computer software]. University Massachusetts, Amherst.

Cohen, A. L. (2012) Fix Align [Computer software]. University of Massachusetts, Amherst.

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the

      lexical decision task. *Quarterly Journal of Experimental Psychology, 31*(3), 489-507.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route

      and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589-600.

Cottrell, N. B. Social facilitation. In C. McClintock (Ed.), *Experimental Social Psychology*.

      London: Holt, Rinehart, & Winston, 1972.

Emerson, M., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task

      investigation. *Journal of Memory and Language, 48*, 1480168.

Filik, R., & Barber, E. (2011). Inner speech during silent reading reflects the reader's regional

      accent. *PLoS ONE, 6*(10).

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research, 27*, 285-319.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings*

      of the North East Linguistics Society (Vol. 32, pp. 112-132). Amherst, Massachusetts:

      GSLA.

Foss, D. J., & Hakes, D. T. (1978). *Psycholinguistics*. Englewood Cliffs, NJ: Prentice-Hall.

Gathercole, S., Baddeley, A.D. (1993). Phonological working memory: A critical building block

      for reading development and vocabulary acquisition? *European Journal of Psychology of*

      Education, 8, 259.

Heavey, C. L., & Hurlburt, R. T. (2008). The phonomena of inner experience. *Consciousness and Cognition, 17*(3), 798-810.

Huey, E. B. (1908/1968). *The psychology and pedagogy of reading.* Cambridge, MA: The MIT Press.

Klinger, E., & Cox, W.M. (1987-1988). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality*, 7, 105-128.

Korba, R. J. (1990, December 1). The Rate of Inner Speech. *Perceptual and Motor Skills, 71*(3), 1043-1052.

Kosslyn, S. M., & Matt, A. M. (1977). If you speak slowly, do people read your prose slowly? Person-partiular speech recoding during reading. *Bulletin of the Psychonomic Society, 9*, 250-252.

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness & Cognition, 20,* 1586-1593.

McGuire, P., Silbersweig, D., Murray, R., David, A., Frackowiak, R., & Frith, C. (1996). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine*, 29-38.

Netsell, R. & Bakker, K. (2017). Fluent and dysfluent inner speech of persons who stutter. Unpublished manuscript. Missouri State University, Springfield, MO.

Newton, A., & de Villiers, J. (2007). Thinking while talking: Adults fail nonverbal false-belif reasoning. *Psychological Science, 18*, 574-579.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 564-573.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in

    integrating information across saccades in word identification and reading. *Journal of*

    Experimental Psychology: Human Perception and Performance, 18(1), 148-162.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-

    Hall.

Rayner, K., & Pollatsek, A. (1994). *The Psychology of Reading.* London: Routledge.

Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are

    automatically activated during reading: Evidence from an eye movement paradigm.

    Psychological Science, 6, 26-32.

Seitz, K. & Schumann-Hengsteller, R. (2002). Phonological loop and central executive processes

    in mental addition and multiplication. *Psychologische Beiträge, 44*(2), 275-302.

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of*

    Experimental Psychology: Human Perception and Performance, 18, 173-184.

Slowiaczek, M. L., & Clifton, C. Jr. (1980). Subvocalization and reading for meaning. *Journal of*

    Verbal Learning & Verbal Behavior, 19(5), 573-582.

Smith, S., Brown, H., Toman, J., & Goodman, L. (1947). The lack of cerebral effects of d-

    Tubocurarine. *Anesthesiology, 8*(1), 1-14.

Sokolov, A. N. (1972). *Inner Speech and Thought.* Springer.

Staub, A., Clifton, C., & Harris, J. (2015). RoboDoc [Computer software]. University of

    Massachusetts, Amherst.

Vilhauer, R. P. (2016). reading voices: An overlooked form of inner speech. *Psychosis, 8*(1), 37-

    47.

Vygotsky, L. (1934/1978). *Mind in society: The development of higher psychological processes.*

(M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.) Cambridge,

Massachusetts, USA: Harvard University Press.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158-

177.

7   Appendixes

Appendix A: Materials (Breen & Clifton, 2011)

1a. Non-reduced/SW/consistent
I know of an old man named Herbert,
Who's known around town as a pervert.
He searches the shops,
For girls in tight tops,
And stands there just too close for comfort.

1d. Non-reduced/WS/inconsistent
I know of an old man named Herbert
Who always the truth tries to pervert.
I know when he lies,
It's all in the eyes.
At hiding it he is no expert.

2a. Non-reduced/SW/consistent
The crew worked so hard for their paychecks
They thought they'd develop a complex
They went out for beer
In a bar by the pier
Then home to their wives in the projects.

2d. Non-reduced/WS/inconsistent
The crew worked so hard for their paychecks
Their work was so terribly complex
They went out for beer
In a bar by the pier
Then home to their wives in the projects.

3a. Non-reduced/SW/consistent
That man applies way too much hair grease.
A friend should suggest a big decrease.
If he's at the beach,
The oil will leach.
He soon will be hearing from Greenpeace.

3d. Non-reduced/WS/inconsistent
That man applies way too much hair grease.
I think the amount he should decrease.
If he's at the beach,
the oil will leach.
He soon will be hearing from Greenpeace.

4a. Non-reduced/SW/consistent
The Soviet spy is a suspect.
The case has but one major defect.
His girlfriend will swear
That he was not there.
But other than that it is perfect.

4d. Non-reduced/WS/inconsistent
The Soviet spy is a suspect.
I heard that he's planning to defect.
He closed his accounts,
Withdrew large amounts,
Adopting a mistrustful affect.

5a. Non-reduced/SW/consistent
We once had a tiresome house guest,
Who loved to read Birdwatcher's Digest.
We teased and we joked,
He was not provoked.
He still thinks that birds are the greatest.

5d. Non-reduced/WS/inconsistent
We once had a tiresome house guest,
Whose humor was painful to digest.
He thinks he's a riot
But we have kept quiet
He still thinks his jokes are the greatest.

6a. Non-reduced/SW/consistent
The gymnast requested a recount
Her score, she thought, rated no discount.
The Austrian judge,
He refused to budge.
He said she had bungled her dismount.

6d. Non-reduced/WS/inconsistent
The gymnast requested a recount
She thought it was wrongful to discount
Her score, but the judge,
He refused to budge.
He said she had bungled her dismount.

7a. Non-reduced/SW/consistent
He tried not to get badly sidetracked.
He needed some raspberry extract.
But in the big shop,
'Twas nary a drop.
His fancy dessert plans were highjacked.

7d. Non-reduced/WS/inconsistent
He tried not to get badly sidetracked
Some essence he wanted to extract
But try as he might,
No tactic worked right
His fancy dessert plans were highjacked.

8a. Non-reduced/SW/consistent
The city must safeguard the seaports,
To save us from dangerous imports.
With so many ships,
There can be no slips.
The tankers will all now need escorts.

8d. Non-reduced/WS/inconsistent
The city must safeguard the seaports,
Because of how much it now imports.

With so many ships,
There can be no slips.
The tankers will all now need escorts.

9a. Non-reduced/SW/consistent
The man who asked you for a consult
Was given a horrible insult.
So this is the end,
He won't ask again,
It seems that you can't be an adult.

9d. Non-reduced/WS/inconsistent
The man who asked you for a consult
Is no-one you wanted to insult.
So this is the end,
He won't ask again,
It seems that you can't be an adult.

10a. Non-reduced/SW/consistent
The teacher assigned them a project
To find an unusual object.
The parents said "Wait, she's not thinking
straight."
The children think she is just perfect.

10d. Non-reduced/WS/inconsistent
The teacher assigned them a project that forced
many parents to object.
The parents said Wait, she's not thinking
straight.
the children think she is just perfect.

11a. Non-reduced/SW/consistent
There once was an old man named Kermit,
Who hunted without any permit.
He wouldn't obey,
They sent him away.
Now he lives alone as a hermit.

11d. Non-reduced/WS/inconsistent
There once was an old man named Kermit
Whose gambling his wife would not permit.
He wouldn't obey,
She sent him away.
Now he lives alone as a hermit.

12a. Non-reduced/SW/consistent
He couldn't hide all of his misdeeds,
But made off with all of the proceeds.
From charity giving,
He has made a living.
Some jail time is what this guy needs.

12d. Non-reduced/WS/inconsistent
He couldn't hide all of his misdeeds
On Monday his retrial proceeds.
I hope that the jury
Convicts in a hurry.
And I'd like to give him a nosebleed.

13a. Non-reduced/SW/consistent
There once was a young man named Ernest,
Who sponsored a violent protest.
When they asked him why,
He said, with a sigh,
"I wanted to open a wasp's nest."

13d. Non-reduced/WS/inconsistent
There once was a young man named Ernest,
Who rounded up people to protest.
When they asked him why,
He said, with a sigh,
"I wanted to open a wasp's nest."

14a. Non-reduced/SW/consistent
That basketball star's like a bloodhound.
He seeks out and catches each rebound.
He always plays tough,

But never too rough,
Although it's just hoops on the playground.

14d. Non-reduced/WS/inconsistent
That basketball star's like a bloodhound. he
waits for each jumpshot to rebound.
He always plays tough,
But never too rough,
Although it's just hoops on the playground.

15a. Non-reduced/SW/consistent
I met an old friend who played baseball,
Who warned of a new safety recall
With all of the noise
About lead in toys,
We're all better off playing stickball.

15d. Non-reduced/WS/inconsistent
I met an old friend who played baseball,
But what his name was I can't recall.
He told me "No worry,"
And left in hurry.
He said he was late for a phone call.

16a. Non-reduced/SW/consistent
Last year I created a stock fund.
And managed to get a big refund.
But now there's no question,
We're in a recession.
Investors are all being cautioned.

16d. Non-reduced/WS/inconsistent
Last year I created a stock fund.
The fees they would happily refund
But now there's no question,
We're in a recession
Investors are all being cautioned.

17a. Non-reduced/SW/consistent
I read an unusual essay
'Bout how they conducted a survey.
The polsters, how rude,
Showed up in the nude.
I burned it right up in the ashtray.

17d. Non-reduced/WS/inconsistent
I read an unusual essay
Describing how folks tried to survey.
The polsters, how rude,
Showed up in the nude.
I burned it right up in the ashtray.

18a. Non-reduced/SW/consistent
The cops have a negative affect
Concerning their most recent suspect.
Stay out of their hair!
They won't treat you fair.
Their treatment may well leave you abject.

18d. Non-reduced/WS/inconsistent
The cops have a negative affect
Toward people they manage to suspect.
Stay out of their hair!
They won't treat you fair.
Their treatment may well leave you abject.

19a. Non-reduced/SW/consistent
A striking young woman named Rembrandt,
From Portugal, she was a transplant.
She hated our food,
And found us quite rude.
Her choice to live here she did recant.

19d. Non-reduced/WS/inconsistent
A striking young woman named Rembrandt
Had roses she wanted to transplant.
They withered and died.

She cried, and she cried.
"Come back to me, my little houseplant."

20a. Non-reduced/SW/consistent
To get to the local gym's squash court,
You must take municipal transport
It will take some time,
And is hard to find.
I'm happy to serve as your escort.

20d. Non-reduced/WS/inconsistent
To get to the local gym's squash court,
Your gear should be ready to transport.
It will take some time,
And is hard to find.
I'm happy to serve as your escort.

21a. Reduced/SW/consistent
You must hear my story, your highness.
I have the young princess's address.
Arrive there at two,
She's waiting for you.
Just be sure to treat her with kindness.

21d. Reduced/WS/inconsistent
You must hear my story, your highness.
Your habits I find I must address.
It's not a good thing,
When our future king,
Behaves like a kiddie at recess.

22a. Reduced/SW/consistent
The guy who got lost on a flyby
Dropped all of his bombs on an ally.
For him it was tragic,
And not the least magic.
The enemy cheered at the bull's eye.

22d. Reduced/WS/inconsistent
The guy who got lost on a flyby
Killed folks with whom we want to ally.
For them it was tragic,
And not the least magic.
Our friends shouldn't be in the bull's eye.

23a. Reduced/SW/consistent
I just saw a dog and a tomcat,
Engaged in some furious combat,
But I was surprised,
And then hypnotized,
When into the fray walked a wombat.

23d. Reduced/WS/inconsistent
I just saw a dog and a tomcat,
That we must be ready to combat,
Or else they will bite,
With all of their might,
Our sweet little domestic wombat.

24a. Reduced/SW/consistent
I heard someone say through the grapevine:
The farmer is driving his combine
To harvest the yields
Of all of his fields,
And have them all shipped on the rail line.

24d. Reduced/WS/inconsistent
I heard someone say through the grapevine:
That farmer is hoping to combine
His own garden's yields,
With those from my fields,
And have them all shipped on the rail line.

25a. Reduced/SW/consistent
I processed some prints in the darkroom
Of people I'd met on a commune.
They pray to the sun,

And have lots of fun,
And frolic like kids on the sand dunes.

25d. Reduced/WS/inconsistent
I processed some prints in the darkroom
Of folks who just wanted to commune
On clothing-free seashores,
The Caymans, or Azores,
And frolic like kids on the sand dunes.

26a. Reduced/SW/consistent
If out in the mountains you backpack,
Your team must agree to this compact:
Don't cut down the trees,
Or bother the bees.
You don't want to make a big impact.

26d. Reduced/WS/inconsistent
If out in the mountains you backpack,
Your gear must be basic and compact.
Don't cut down the trees,
Or bother the bees.
You don't want to make a big impact.

27a. Reduced/SW/consistent
We stayed in the wooods at a campground,
Which wasn't too far from a compound
Where there was a crowd
That drank and was loud.
Their rudeness our party did dumbfound.

27d. Reduced/WS/inconsistent
We got that old dog at the pound
Our sadness will surely compound
For now he is sick,
Can do no more tricks,
And soon will be put in the ground.

28a. Reduced/SW/consistent

There was a young heroin addict,
Who ended up causing a conflict.
His auto was stopped,
And he shot a cop,
And now he's a federal convict.

28d. Reduced/WS/inconsistent
There was a young heroin addict,
Whose habits and others did conflict.
His auto was stopped,
And he shot a cop,
And now he's a federal convict.

29a. Reduced/SW/consistent
The athlete who just failed a drugtest,
Will soon face a challenging contest.
His one last resort,
To take it to court,
By filing a certified protest.

31a. Reduced/SW/consistent
In nothing but jeans and a t-shirt,
That man took a trip 'cross the desert.
He soon was quite beat,
Succumbed to the heat,
He wasn't a survival expert.

31d. Reduced/WS/inconsistent
In nothing but jeans and a t-shirt,
A soldier his squad chose to desert.
He soon was quite beat,
Succumbed to the heat,
He wasn't a survival expert.

32a. Reduced/SW/consistent
I know of an elegant female
Her outfits lack no fashion detail.

She gave me advice:
I don't pay full price.
I buy all my clothing at wholesale.

32d. Reduced/WS/inconsistent
There once was a woman named Gail
Who wanted her car to detail.
A nice guy said he
Would do it for free,
And later they went out for ale.

29d. Reduced/WS/inconsistent
The athlete who just failed a drugtest,
Is planning the charges to contest.
His one last resort,
To take it to court,
By filing a certified protest.

30a. Reduced/SW/consistent
Although that young man is an addict,
He really should not be a convict.
He needs some support,
It's better than court,
And jail time will cause too much conflict.

30d. Reduced/WS/inconsistent
Although that young man is an addict,
I think that the judge shouldn't convict.
He needs some support,
It's better than court,
And jail time will cause too much conflict.

33a. Reduced/SW/consistent
There was a young woman whose nude dance
Made everyone flock to the entrance.
The men it drove wild,
Whenever she smiled,
But she says she's not seeking romance.

33d. Reduced/WS/inconsistent
There was a young woman whose nude dance
Would always the gentlemen entrance.
She drove them just wild,
Whenever she smiled,
But she says she's not seeking romance.

34a. Reduced/SW/consistent
There once was a penniless peasant,
Who couldn't afford a nice present.
For so little yield,
He worked in the field,
And work there was really unpleasant.

34d. Reduced/WS/inconsistent
There once was a penniless peasant,
Who went to his master to present
His gripes and complaints.
He showed no restraint.
The outcome was rather unpleasant.

35a. Reduced/SW/consistent
There once was a crusty old recluse,
Who grew the most wonderful produce.
His lemons and limes
Were truly sublime,
And made the most wonderful fruit juice.

35d. Reduced/WS/inconsistent
There once was a crusty old recluse,
Whose garden great harvests would produce.
His lemons and limes
Were truly sublime,
And made the most wonderful fruit juice.

36a. Reduced/SW/consistent
With all of their time spent at recess,
The children make no forward progress.

They won't be succeeding,
In writing and reading,
While working on physical fitness.

36d. Reduced/WS/inconsistent
With all of their time spent at recess,
The children will soon fail to progress.
They won't be succeeding,
In writing and reading,
While working on physical fitness.

37a. Reduced/SW/consistent
I noticed a ruinous defect
In one of the candidate's projects.
If he can't explain,
He'll garner disdain.
The voters will find the man suspect.

37d. Reduced/WS/inconsistent
I noticed a ruinous defect
In what that new candidate projects.
He oftens looks dour,
And out-and-out sour.
The voters are finding him suspect.

38a. Reduced/SW/consistent
In a voice that was piercing and treble,
The serfs were inspired by a rebel.
Then facing the troops,
In small tight-knit groups,
They let fly a volley of pebbles.

38d. Reduced/WS/inconsistent
In a voice that was piercing and treble,
The leader urged peasants to rebel.
Then facing the troops,
In small tight-knit groups,
They let fly a volley of pebbles.

39a. Reduced/SW/consistent
There once was a young man named Eckerd
Who broke an old pole-vaulting record.
His star rose so fast,
But fame would not last.
It turned out his past was quite checkered.

40a. Reduced/SW/consistent
The judges must all watch the replay
To find out which team won the relay.
To my unskilled eye,
It looked like a tie.
The runners all wait on the raceway.

39d. Reduced/WS/inconsistent
There once was a young man named Eckerd
Whose pole-vaulting feats they did record
His star rose so fast,
But fame would not last.
It turned out his past was quite checkered.

40d. Reduced/WS/inconsistent
The judges must all watch the replay.
Results to the coach they will relay
To my unskilled eye,
It looked like a tie.
The runners all wait on the raceway

Appendix B: Exit survey and VISQ

Please complete this short 2-page questionnaire about your preferences and behaviors.

What did you notice about the experiment that was interesting or odd?

What was your strategy for reading the limericks?

What do you think we were studying in the experiment?

What do you think our hypothesis was?

During the experiment, did you use any strategies to read the sentences? Please explain.

Participant experimental ID

Birth date

Age in years

Dominant hand

Native language(s)

Age (in years) at which you began speaking in English. If you are a native speaker of English, write "birth."

Sex

EHNICITY—Do you consider yourself to be Hispanic or Latino? (See definition below.) Select one.

- Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish origin, regardless of race. The term "Spanish origin" can be used in addition to "Hispanic or Latino".
- Not Hispanic or Latino
- Check here if you do not wish to provide this information.

RACE—What race do you consider yourself to be? Select one or more of the following.
- American Indian or Alaska Native. A person having origins in any of the original peoples of North, Central, or South America, and who maintains tribal affiliation or community attachment.
- Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including for example Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand and Vietnam.
- Black or African American. A person having origins in any of the black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "Black" or "African American".
- Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
- White. A person having origins in any of the original peoples of Europe, the Middle East or North Africa.
- Check here if you do not wish to provide this information.

Are you currently in PSYC 001?

Please list any other Psychology courses you have taken or are currently taking.

Please list any other Linguistics courses you have taken or are currently taking.

Do you have any known hearing problems? Type "Yes" or "No." If Yes, please explain.

Do you have one of the following: normal vision, or vision corrected to normal by contacts or glasses?

Have you had any known brain injuries? Type "Yes" or "No." If Yes, please explain.

Do you have any known hearing or auditory impairments? Type "Yes" or "No." If Yes, please explain.

Do you have any known auditory processing disorder? Type "Yes" or "No." If Yes, please explain.

1. I think to myself in words using brief phrases and single words rather than full sentences.
2. When I am talking to myself about things in my mind, it is like I am going back and forward asking myself questions and then answering them.
3. I hear the voice of another person in my head. For example, when I have done something foolish I hear my mother's voice.
4. I experience the voices of other people asking me questions in my head.
5. I hear other people's voices nagging me in my head.
6. My thinking in words is more like a dialogue with myself, rather than my own thoughts in a monologue.
7. I think to myself in words using full sentences.
8. My thinking to myself in words is like shorthand notes, rather than full, proper, grammatical English.

9. I think in inner speech about what I have done, and whether it was right or not.

10. When I am talking to myself about things in my mind, it is like I am having a conversation with myself.

11. I talk silently to myself telling myself to do things.

12. I hear other people's actual voices in my head, saying things that they have never said to me before.

13. I talk back and forward to myself in my mind about things.

14. My thinking in words is shortened compared to my normal out-loud speech. For example, rather than saying to myself things like 'I need to go to the shops', I will just say 'shops' to myself in my head.

15. If I were to write down my thoughts on paper, they would read like a normal grammatical sentence.

16. I hear other people's actual voices in my head, saying things that they actually once said to me.

17. I talk silently to myself telling myself not to do things.

18. I evaluate my behavior using my inner speech. For example I say to myself, 'that was good' or 'that was stupid.'

Appendix C: Additional resources

Additional resources and relevant materials can be found at https://github.com/itang1/limerick1.