High-level listening comprehension in advanced English as a second language: Effects of the first language and inhibitory control

Mandy Wigdorowitz[a,b] ⓘ, Ana I. Pérez[c] ⓘ and Ianthi M. Tsimpli[a] ⓘ

[a]Theoretical and Applied Linguistics, University of Cambridge, Cambridge, United Kingdom; [b]Department of Psychology, University of Johannesburg, Johannesburg, South Africa and [c]Department of Experimental Psychology, University of Granada, Granada, Spain

Abstract

English is imposed as the language of instruction in multiple linguistically diverse societies where there is more than one official language. This might have negative educational consequences for people whose first language (L1) is not English. To investigate this, 47 South Africans with advanced English proficiency but different L1s (L1-English vs. L1-Zulu) were evaluated in their listening comprehension ability. Specifically, participants listened to narrative texts in English which prompted an initial inference followed by a sentence containing an expected inference or an unexpected but plausible concept, assessing comprehension monitoring. A final question containing congruent or incongruent information in relation to the text information followed, assessing the revision process. L1-English participants were more efficient at monitoring and revising their listening comprehension. Furthermore, individual differences in inhibitory control were associated with differences in revision. Results show that participants' L1 appears to supersede their advanced English proficiency on highly complex listening comprehension.

Introduction

The ability to draw inferences from auditory and written input is crucial for comprehension and successful educational outcomes and is especially relevant in linguistically diverse contexts where learners have heterogeneous language backgrounds but are educated in the predominant, privileged language of the country. Such a case is South Africa, where tertiary education is almost exclusively received through the medium of English (Madadzhe, 2019), though it is not the first language (L1) for the majority of the population but for only 9.6% of citizens (Statistics South Africa, 2012). This makes South Africa a unique context to investigate language effects in listening comprehension in a sample that scarcely (if at all) comprises 'pure' monolinguals but is home to individuals that widely differ in their L1 and degree of multilingualism. Accordingly, the main aim of this study is to investigate how advanced English proficient South African adults differing in their L1 (English or Zulu) perform on listening text comprehension in the language of instruction (English). We first provide an overview of the language and educational context in South Africa, followed by an account of the cognitive and linguistic processes involved in text comprehension in both reading and listening modalities.

South African linguistic and educational context

Young adults completing tertiary education in South Africa are exposed to regular comprehension of complex and novel content. Learners entering a university are required to, at minimum, understand, speak, read, and write in standard, university-level English if they are to be adept and competitive (South African Department of Education, 2008). While this seems like an unreasonable expectation given how few South Africans have English as an L1, the post-apartheid government has attempted to put educational safeguards in place that aim to ensure adequate English competency by tertiary education, while still promoting multilingualism (Constitution of the Republic of South Africa, 1996; Department of Education, 1997). While, in principle, the language-in-education policy is linguistically inclusive, in reality, multilingual education still remains imbalanced, with English (and to a lesser extent, Afrikaans) being prioritised and disproportionally supported (Drummond, 2016; van Wyk, 2014). English medium schools are by far the most prevalent across South Africa (Webb, Lafon & Pare, 2010), so over the past two decades the majority of South Africans have received education in English as the language of learning and teaching throughout their schooling career. This disparity is multidimensional but in a large part due to the elevated status that English occupies, as well as access to educational resources and teaching

material in English, and the conception that English is a global *lingua franca* that can lead to more long-term social and economic opportunities and mobility (Madadzhe, 2019; Mesthrie, 2002). Similarly, English is ubiquitous across government, media, and industry, so it is indeed privileged in this regard and considered the country's linguistic bridge.

It can therefore be assumed that South Africans entering university have prolonged and sustained exposure to English both in and out of the school context as well as high levels of proficiency to meet the language requirements necessary to be accepted to an institution of higher education.[1] Few studies have investigated differences between highly English proficient adult learners, but who enter educational systems with different L1s. We are particularly interested in L1 as opposed to LINGUALISM STATUS (monolingual, bilingual) because L1 is a highly informative indicator of foundational language background in the South African context, given the (mis)alignment that it has in comparison to the medium of tertiary education. Importantly, the question of whether individuals' L1 supersedes their proficiency in listening text comprehension has not previously been explored. There is, however, concrete evidence to suggest that comprehension ability is predicted by proficiency (Filippi, Leech, Thomas, Green & Dick, 2012), particularly in school-aged children (Babayiğit & Shapiro, 2020; Melby-Lervåg & Lervåg, 2014; Pretorius & Spaull, 2016), but it is unclear whether young adults situated in a multilingual context who have advanced proficiency in English (i.e., the language of learning and teaching) but differ in their L1 (in this case, L1-Zulu) are able to perform equally well on a demanding comprehension task compared to L1-English participants.

High-level text comprehension processes

Comprehension is the process of extracting meaning from visual and/or auditory stimuli as opposed to a mere verbatim record of the presented content. It involves the interplay between world-knowledge (top-down processing) in conjunction with the accumulated content-based information of the text (bottom-up processing) so that the comprehenders can construct an integrated and coherent mental representation of the content that clarifies its meaning (for reviews, see Aryadoust, 2019; McNamara & Magliano, 2009). This mental representation is referred to as a SITUATION MODEL (Kintsch, 1998). As each word or phrase is processed, it is actively integrated into the situation model of comprehenders such that text coherence can be achieved. At the lexical or sentence level, word meanings are retrieved and systematically grouped into meaningful grammatical units, while at the discourse level, other higher-level cognitive skills are engaged to be able to construct a unified, coherent, and accurate representation of the content in real time (Rapp, van den Broek, McMaster, Kendeou & Espin, 2007; van Dijk & Kintsch, 1983). Relevant higher-level comprehension processes that play a role in text coherence include INFERENCING, MONITORING, and REVISION.

INFERENCING is the ability to deduce information from content without it being explicitly referred to (Cook & O'Brien, 2017; Graesser, Singer & Trabasso, 1994; Zwaan & Radvansky, 1998). An important inference affecting language processing is prediction – the capacity to precedingly anticipate incoming information or adapt to situations where conflicting or ambiguous

information arises that deviates from one's expectations (Bubic, von Cramon & Schubotz, 2010; Pickering & Gambi, 2018). For example, in (1) the target word "*son*" is highly semantically constraining and expected, while in (2) the target word "*dog*" is less semantically constraining – that is, unexpected, but still plausible. When presented with examples such as (2) where the inference is not expected, prediction errors (e.g., processing delays, decline in accuracy) are likely to arise, especially where the strength of the generated prediction is robust, or the content is presented in a second language (Kaan & Grüter, 2021). Importantly, prediction errors are not confined to single sentences like (1) and (2) but may emerge in longer narratives where knowledge-based inferences have been generated based on the establishing content, as in (3).

(1) The mother has two daughters and one *son*
(2) The mother has two daughters and one *dog* (adapted from Dijkstra, Van Hell & Brenders, 2015)

COMPREHENSION MONITORING is the active, metacognitive strategy that facilitates awareness of what is being comprehended, whereby efficient comprehenders are able to rapidly notice information that does not align with previously presented content so as to disambiguate their predictions (Baker, 1989; McNamara, 2007; Wagoner, 1983; Zhang, 2017). A consistent outcome is a time cost on unexpected compared to expected sentences because of noticeable disruptions to text coherence (O'Brien, Rizzella, Albrecht and Halleran, 1998; Rinck & Weber, 2003). For instance, following passage (3), participants have to adjust their expectations with reference to additional incoming content that is either expected and aligns (4a) or unexpected and misaligns (4b) with the prior information. Efficient comprehenders are able to rapidly notice information that does not align with previously presented content in order to disambiguate their predictions (McNamara, 2007; Zhang, 2017).

(3) In the lake by my house there are many animals that make quacking noises. Yesterday the sounds were really loud, so I went to see which animals were making them.
(4a) On a big rock in the middle of the lake, I could see the *ducks*
(4b) On a big rock in the middle of the lake, I could see the *turtles*

Moreover, in situations where upcoming information does not match with a previously generated inference as the narrative unfolds and further information is presented, comprehenders need to review their existing interpretation and update their situation model so that the new and unexpected content becomes meaningfully integrated into the discourse context, in a process known as REVISION (Rapp & Kendeou, 2007). Participants need to update their situation model in order to identify whether the content is congruent (4a → 5a or 4b → 5b) or incongruent (4a → 5b or 4b → 5a) in relation to preceding information. Revision of the situation model is achieved only if comprehension of the text is adequately regulated by inhibiting the original interpretation that was encoded into the situation model in favour of the new, updated interpretation – an updating process that is both linguistically and cognitively demanding (Kendeou, Smith & O'Brien, 2013; Rapp & Kendeou, 2007). Central to the revision process is a distinction between the processes of updating and outdating, whereby the former involves the encoding and integration of new information into the mental representation of the text, while the latter entails a disregarding and replacement of the no

---

[1]See, for example https://www.wits.ac.za/internationalstudents/requirements/english-language-proficiency/ (last accessed 19/01/2023).

longer-true, irrelevant, and outdated information that disrupts the text's coherence (Kendeou et al., 2013). Outdated information is expected to decay in line with the presentation of new, harmonious information and should be deactivated from the current mental model, losing its accessibility (Zwaan & Radvansky, 1998). Much of the time this is borne out such that comprehenders can successfully update their situation model (de Vega, 1995; Rapp & Taylor, 2004). However, it has also been demonstrated that while revision can be achieved, processing delays may be incurred as a direct consequence of maintaining the outdated information in the situation model (de Vega, Urrutia & Riffo, 2007; Guéraud, Harmon & Peracchi, 2005; O'Brien, Cook & Guéraud, 2010); Rapp & Kendeou, 2007). In such cases, comprehension can be disrupted.

(5a) There were animals with brown and green *feathers* making noise
(5b) There were animals with brown and green *shells* making noise (adapted from Pérez, Joseph, Bajo & Nation, 2016)

The majority of research conducted on high-level comprehension processes has focussed on reading rather than listening comprehension (for a review, see McNamara & Magliano, 2009) and even fewer studies have evaluated comprehension from a multimodal perspective (Hu & Jiang, 2011; Manfredi, Cohn, Andreoli & Boggio, 2018; Pérez, Schmidt, Kourtzi & Tsimpli, 2020; Wu & Ma., 2016). Despite this asymmetry, the processes involved in reading and listening comprehension appear to overlap in critical ways and rely on similar bottom-up and top-down mechanisms (Rost, 2005, 2016). They also act simultaneously and in parallel by drawing on various types of knowledge sources including linguistic, world, and context-based knowledge (Vandergrift, 2007) as well as depend on the proficiency/language background of the listener, their working memory capacity, and the amount of inhibitory control needed to process the content (Kim, 2015; Pérez, Paolieri, Macizo & Bajo, 2014; Rapp et al., 2007). In fact, although for proficient listeners the process can be largely automatic, listeners have some control over their degree of engagement and attentional allocation of the input being attended to (Bodie, Worthington, Imhof & Cooper, 2008). That is, listening can be effortful with more challenging contexts (e.g., L2 comprehension) compared to less demanding contexts (e.g., L1 comprehension). Crucially, there is a gap in the literature of listening comprehension when the language tested is not the participant's L1 and when multilingualism is characteristic of the individual and societal context.

High-level text comprehension and inhibitory control

Inhibitory control (IC) and working memory have routinely been associated with comprehension ability (Pérez et al., 2014; Tarchi, Ruffini & Pecini, 2021). In a significant study for the present research, Pérez et al. (2016) investigated how adults monitor and revise their situation model after encountering unexpected information. Using the mismatch detection paradigm, participants were presented with a series of short narrative texts where the first sentences primed a knowledge-based inference as in (3), followed by a sentence that either brought an expected ("ducks") or unexpected but plausible ("turtles") word, as in (4a) and (4b). Results revealed that readers were able to detect a mismatch on unexpected compared to expected words, indicating a good ability to monitor inferential information in real time. In

addition, following the establishing text, participants were then presented with one of two comprehension sentences that brought either congruent or incongruent information in reference to the content of the prior text, as in (5a) and (5b). This time, readers with lower verbal (but not visuospatial) working memory had longer reading times and made more regressions on congruent sentences only after being exposed to the unexpected word ("turtles" → "shells"). The authors suggested that the revision process is more cognitively demanding than the monitoring process and it depends on (verbal) working memory capacity (also see Carretti, Borella, Cornoldi & De Beni, 2009). Importantly, a similar relationship has been found with IC and its influence on the quality of the mental representation that is constructed during reading and oral single word comprehension (Arrington, Kulesz, Francis, Fletcher & Barnes, 2014; Blumenfeld & Marian, 2011). Interestingly, the effect of working memory in revision disappears when IC is taken into account to explain individual differences in multimodal (auditory-verbal and visual-pictorial) comprehension in L1-English speakers (Pérez et al., 2020) since IC is thought to encompass aspects of working memory (Tiego, Testa, Bellgrove, Pantelis & Whittle, 2018). IC is therefore an important executive function to consider when measuring the process of revision in text comprehension, but it has not been solely assessed during listening comprehension.

Finally, a study investigating high-level cognitive processes during reading comprehension in the L1 and L2 also found that executive control (in addition to L2 proficiency) explained individual differences in the revision process (Pérez, Hansen & Bajo, 2019). Specifically, efficient revision in the L1 was related to a balance between proactive control (the ability to sustain information related to a meaning goal in anticipation of a cue) and reactive control (the ability to reactivate meaning goals as a response to a cue), whereas better revision in the L2 was associated with strong proactive control. Thus, a relevant matter here is to understand whether in developing and multilingual countries such as South Africa, where English is imposed as the official language in tertiary education, L1-English speakers have better performance on high-level listening comprehension ability, or whether people with a different L1 background (L1-Zulu) but high L2-English[2] proficiency perform equally well.

The present study

The main aim of this study was to investigate how advanced English proficient South African adults differing in their L1 (English or Zulu) perform on listening text comprehension in the language of instruction (English). The present study expands on research that has assessed inferential monitoring and revision during reading and multimodal comprehension (Pérez et al., 2016, 2020) by presenting narrative texts only in the auditory modality. By comparing L1-English vs. L1-Zulu populations in their ability to perform inferencing, monitoring, and revision, we can understand the importance of L1 experience over and above task proficiency in individuals from multilingual contexts. An additional aim was to understand if the effects derived from the main goal were explained by individual differences in IC, especially regarding the revision process.

To test our aims, we used an auditory version of the mismatch detection task (Pérez et al., 2016) – henceforth, AMDT. In this

---

[2]While L2 does not always strictly refer to participants' second language, we use the term to denote a language acquired after the L1.

task, narrative texts comprising three parts (establishing, critical, and comprehension) are auditorily presented to participants in succession (see Table 1). The first two sentences establish the context of the story by describing a scenario that biases a knowledge-based inference generated at the level of the situation model ("ducks"). Response latency (RL) after listening to the establishing sentences provides a "baseline RL". A third sentence is then presented in one of two possible conditions that either aligns and is consistent with the prior primed narrative content (expected condition: "ducks") or misaligns and is unlikely but still provides valid and plausible information (unexpected condition: "turtles"). In this sentence, the unexpected condition requires participants to re-evaluate their mental representation in relation to their initial inference, while in the expected condition no re-evaluation is necessary since the content confirms the prediction biased in the establishing context. Only the last word of the critical sentence is manipulated for each narrative with the rest of the wording being identical, thus allowing for participants to experience either expected or unexpected information at the same point in the critical sentence in all cases. RL after listening to the critical sentences gives the "target word RL". Importantly, to extract a more pure processing time, the target word RL is divided by the baseline RL (i.e., "pure target word RL"), which is used as an index of comprehension monitoring.

Lastly, a final comprehension sentence is again auditorily presented also in one of two possible conditions: a congruent sentence in which the content relates to the information from the target concept in the critical sentence (e.g., "feathers" for "ducks" or "shells" for "turtles"); or an incongruent sentence where the content does not relate to the target concept in the critical sentence (e.g., "shells" for "ducks" or "feathers" for "turtles"). Here, participants have to revise the content presented to them and indicate whether the comprehension sentence matches the information presented in the critical sentence. RL from the onset of the question mark to the onset of the "Yes/No" response to the comprehension sentence gives the "sentence RL". Similar to the monitoring process, this measure is divided by the baseline RL to obtain a more pure processing time (i.e., "pure sentence RL"), and this measure works as an index of revision. Finally, accuracy scores are also recorded in the comprehension sentence, as a second measure of revision.

Aims and hypotheses

The first aim of this study is to investigate the effect of L1 on high-level comprehension processes in text listening comprehension in English proficient young South African adults completing a tertiary education, and secondly, to evaluate whether individual differences in IC modulate prior effects. Accordingly, our hypotheses are as follows. For the COMPREHENSION MONITORING PROCESS, we expect greater prediction interference (increased pure target word RLs) for all comprehenders when they are presented with unexpected compared to expected critical words (Hypothesis 1). Furthermore, L1-Zulu participants are predicted to respond comparatively slower to the unexpected condition than L1-English participants because of the misalignment between their L1 and that of the task, while no differences between groups are anticipated for the expected condition where no conflict arises (Hypothesis 2).

For the REVISION PROCESS, given that RLs have been found to increase and accuracy to decrease when sentence coherence is disrupted in general and as a consequence of proficiency (e.g., Pérez et al., 2016, 2019, 2020), we expect pure sentence RLs to

be higher (Hypothesis 3) and accuracy to be lower (Hypothesis 4) for comprehenders in conditions where there is prediction interference (expected-incongruent, unexpected-congruent, unexpected-incongruent), specifically where revision is needed, i.e., for critical sentences containing unexpected target words followed by congruent or incongruent comprehension sentences. Similarly, we predict RLs to be higher (Hypothesis 5) and accuracy to be lower (Hypothesis 6) in L1-Zulu participants compared to L1-English participants on all but the expected-congruent (easiest) condition, given the mismatch between Zulu participants' L1 and the language of the task, while for L1-English participants, no such interference should emerge. These findings would indicate difficulties for L2 English speakers to revise their situation model during high-level listening comprehension in English.

According to the second aim, IC is not expected to explain individual differences for young adults in the monitoring process, because as compared to the revision process, text monitoring is easier and less cognitively demanding (Pérez et al., 2016; Hypothesis 7). However, for the revision process, we predict IC to modulate the effect of listening comprehension performance given that comprehenders with higher IC could be better at managing this type of conflict by suppressing prior conflicting information (Pérez et al., 2020), and in particular, we expect IC to be associated with conditions that tap into the revision process specifically (unexpected-congruent and unexpected-incongruent conditions), such that those with better IC will answer the comprehension sentences more efficiently and/or with lower RLs (Hypothesis 8). Finally, L1-Zulu participants with higher IC could also show a more comparable pattern in the revision process, similar to the one manifested by L1-English participants (Hypothesis 9; Pérez et al., 2019).

Method

Participants

Participants were 47 South Africans between 18-24 years old ($M_{age} = 19.53$, $SD = 1.37$) who were completing an undergraduate degree at an English-medium university in Johannesburg (years of tertiary education: $M = 1.11$, $SD = .98$). Based on their reported language background, participants were divided into two groups: L1-Zulu ($n = 22$; $M_{age} = 19.59$, $SD = 1.37$) and L1-English ($n = 25$; $M_{age} = 19.48$, $SD = 1.39$) speakers, where L1 is operationalised as the first language participants learned as a child. Participants also listed their home, spoken, and written languages, the language(s) they were formally taught in during primary and secondary schooling, as well as their use of English among various interlocutors. Descriptive details about this linguistic information can be found in Appendix A. Importantly, all participants were highly proficient in English, but, as expected, L1-Zulu participants used English for fewer years on average than L1-English participants: 14.71 compared to 18.80 years, $t(44) = -5.51$, $p < .001$.

Materials

High-level listening comprehension processes

Text listening comprehension was evaluated by means of the AMDT (see Table 1). Participants were presented with 44 (4 practice, 40 experimental) culturally appropriate audio-recorded

Table 1. Example of a narrative story from the auditory mismatch detection task showing sentence type and narrative information, factors and conditions, cognitive processes, response recording, and dependent variables assessed in each process.

Sentence type and narrative information	Factor and conditions	Cognitive process	Response recording	Dependent variable
Establishing sentences In the lake by my house there are many animals that make quacking noises. Yesterday the sounds were really loud, so I went to see which animals were making them.	Context establishment	*Inferencing* ("*ducks*")	From fixation cross (+) at end of establishing sentence to spacebar press to cue onset of critical sentence	Baseline RL
Critical sentence On a big rock in the middle of the lake, I could see the "**ducks/turtles**".	**Expectancy** Expected = *ducks* Unexpected = *turtles*	*Monitoring*	From fixation cross (+) at end of critical sentence to spacebar press to cue onset of comprehension sentence	Pure target word RL[a]
Comprehension sentence There were animals with brown and green "**feathers/shells**" making noise.	**Congruency** Congruent or Incongruent	*Revision*	From question mark (?) at end of comprehension sentence to answer of Yes/No question	Pure sentence RL[b] and Accuracy

Note. RL = Response latency. [a]Critical sentence RL / baseline RL; [b]Comprehension sentence RL / baseline RL.

four-sentence narrative texts sourced from Pérez et al. (2020)[3], in one of the two possible critical sentence conditions (expected or unexpected) and one of the two possible comprehension sentence conditions (congruent or incongruent). As such, four versions were created and counterbalanced so that each of the 40 narratives were presented to participants only once in one of the four cross-conditions, with 10 trials per cross-condition (expected-congruent, expected-incongruent, unexpected-congruent, unexpected-incongruent). The trials were presented in randomised order across two blocks. To begin each trial, the establishing sentence was played when participants pressed the spacebar. Immediately afterwards, a fixation cross appeared in the centre of the screen. Participants pressed the spacebar again to cue the soundtrack of the critical sentence, and when it was finished, a second fixation cross appeared. Participants pressed the spacebar to hear the final comprehension sentence. This time, after the comprehension sentence was played, a question mark appeared and participants were instructed to press "Yes" ('M' key) if this sentence matched with the information presented in the critical sentence, or "No" ('Z' key) if it did not match, as quickly and as accurately as possible.

As it was mentioned, RLs were recorded at three different points: (1) from the onset of the fixation cross after listening to the establishing sentences to the onset of the spacebar key press (baseline RL); (2) from the onset of the fixation cross after listening to the critical sentence to the onset of the spacebar key press (target word RL); and (3) from the onset of the question mark to the onset of the "'Yes/No" response to the comprehension sentence (sentence RL) (see Table 1). Accordingly, to assess comprehension monitoring, we calculated a "pure target word RL" index by dividing the target word RL by the baseline RL, whereas revision was evaluated by a "pure sentence RL" index, which was extracted by dividing the sentence RL by the baseline RL. These indexes avoided any influence from RL speed. Finally, accuracy

scores were also recorded in the comprehension sentence, as a second measure of revision.

Individual differences measure

Inhibitory control

The flanker task from the Attention Network Task was used to assess the executive function of IC (Fan, McCandliss, Sommer, Raz & Posner, 2002). In this task, five horizontal arrows are presented, and participants need to indicate the direction of the central arrow (left or right). Surrounding arrows can either point in the same (congruent condition) or opposite (incongruent condition) direction compared to the central arrow. Twenty-four practice trials followed by 288 main trials were presented over three blocks (96 trials per block, presented in a counterbalanced order with four cue conditions by three target conditions). Accuracy rate across the task was high and approaching ceiling across all conditions ($M = 96.31\%$; $SD = 3.69$). Incorrect trials (3.69%) and extreme values under 200ms and over 1200ms (2.50%) were removed. Reaction times (in milliseconds) were recorded for each trial and averaged across each condition to calculate the conflict effect (mean RT incongruent - mean RT congruent flanker conditions; $Ms = 699.40$ ($SD = 71.26$) and 579.77 ($SD = 70.18$) for incongruent and congruent conditions, respectively). Higher conflict effect scores mean worse IC (i.e., inhibitory cost), whereas the opposite is true for lower conflict effect scores. The IC index was included in the final model to understand whether individual differences in this measure explained differences in the monitoring and revision processes.

Control measures

To ensure we had a valid presentation of participants' ENGLISH PROFICIENCY, both self-assessed and objective measures were used. The self-assessed measure was evaluated by means of the Contextual Linguistic Profile Questionnaire (CLiP-Q), whereas objective proficiency was evaluated with an expressive vocabulary task, a verbal fluency task, and a reading comprehension task (see descriptive statistics in Table 2).

Contextual linguistic diversity

The CLiP-Q (Wigdorowitz, Pérez & Tsimpli, 2020) was used to assess contextual and individual linguistic diversity. It is a

---

[3]Items can be accessed in the supplementary material of Pérez et al. (2020). They include four practice items (one per condition) and items 1, 2, 3, 5, 7, 11, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 26, 27, 28, 29, 31, 32, 34, 35, 36, 37, 40, 42, 43, 44, 45, 47, 48, 49, 51, 53, 54, 57, 58, 60. Since the original (Pérez et al., 2016) and multimodal (Pérez et al., 2020) studies were both carried out in the United Kingdom, the content of each narrative text was evaluated so that only the 40 texts most culturally appropriate and applicable to the South Africa context were retained. For instance, a text was excluded if it referred to the monarchy or snowfall – concepts South Africans may be less familiar with.

Table 2. Mean differences for participants' background measures by language group.

	Range	L1-English	L1-Zulu	F	p	η^2
Self-assessed English proficiency[a]						
Speaking	6–10	9.00 (1.00)	8.55 (1.01)	2.39	.13	.05
Understanding	6–10	9.12 (1.09)	8.64 (1.00)	2.48	.12	.05
Reading	6–10	9.12 (1.01)	8.73 (1.20)	1.48	.23	.03
Writing	6–10	8.96 (1.17)	8.59 (1.10)	1.23	.27	.03
Objective English proficiency						
Expressive vocabulary[b]	10–17	14.84 (1.18)	13.91 (1.93)	4.11	.05	.08
Verbal fluency	15–42	25.08 (4.48)	23.32 (5.13)	1.58	.22	.03
Reading comprehension[c]	11–29	20.60 (.07)	20.09 (3.07)	.23	.64	.01
Working memory	.71–.96	.86 (.06)	.84 (.07)	1.23	.27	.03
Socio-economic status						
Socio-economic status index	1.5–6.0	3.81 (1.30)	3.09 (1.05)	4.26	* .045	.09
School type[d]	2–6	4.04 (.94)	3.23 (.97)	9.11	** .004	.17

Note. [a]range 1 (*very low*) to 10 (*perfect*); [b]norm scores from the WAIS-IV[SA]; [c]range 0 to 30, [d]range 2 to 6. * = $p < .05$, ** = $p < .01$.

comprehensive language profiling measure that captures participants' demographic information (e.g., age, gender), language background (e.g., L1, L2), use (e.g., home, school), proficiency (e.g., speaking, writing), and socio-economic status (SES), emphasising the importance of sociolinguistic context. It has been validated and used in the South African context (Wigdorowitz, Pérez & Tsimpli, 2022). As part of the CLiP-Q, participants evaluated their English proficiency across the modalities of speaking, understanding, reading, and writing on a Likert scale from 1 (*very low*) to 10 (*perfect*). These scores were used as indicators of self-assessed English proficiency.

Expressive vocabulary

To test expressive vocabulary, we used the South African version of the *WAIS-IV^{SA} Vocabulary subtest* (Wechsler, 2008, 2014). Participants were verbally and visually presented with a series of 30 words that increased in difficulty for which they were instructed to provide oral definitions of each. Each definition is awarded a score of 0, 1, or 2 (except for the first three words, whose maximum scores are 1; range = 0–57). This scoring system allows for easy discrimination between high-level and superficial-level responses to the same item. Scores were summed and converted to age-normed standard scores ($M = 10$, $SD = 3$; range = 0–19), with higher scores indicating better performance.

Verbal Fluency

The verbal fluency task assesses lexical retrieval efficiency, tapping into both proficiency and varying levels of executive control (Friesen, Luo, Luk & Bialystok, 2015). In our version, participants were required to generate exemplars from two semantic categories: living (animals) and non-living (fruits), under a 30-second time restriction per condition. One point was awarded for each new and correct word and no points were given for repeated words or if words did not belong to the specified category (e.g., carrot reported as a fruit). Summed scores for the combined categories were calculated.

Reading comprehension

Reading comprehension was assessed using the Adult Reading Test (ART, second edition; Brooks, Everatt & Fidler, 2016). Participants were presented with one practice and three main passages that increased in length and difficulty based on the Flesh Kincaid grade levels – a readability formula assessing the approximate reading grade level of a text (Flesch, 1949; Kincaid, Fishburne, Rogers & Chissom, 1975). After reading each main passage, participants were asked 10 pre-specified comprehension questions based on fact, memory, or inference to assess differing levels of understanding and retention. Correct answers received a score of 1 with a total comprehension score being the sum of all questions (range = 0–30).[4]

Working memory

Working memory was tested by a 2-back task, which measures the storage and updating of information (Kirchner, 1958; Mackworth, 1959). In this task, participants are required to monitor a sequence of digits to determine whether the present digit is the same as the digit that was presented n-digits back; in this case, two digits back, or 2-back. Of a total of 60 digits presented in the main trial, 20 met the 2-back criteria and required a 'J' key press, while 40 did not, requiring no button press, and were considered 'false alarms'. Before the start of the main trials, participants were presented with 10 practice trials comprising three 'hits' and seven 'false alarms'. A composite A' (A-prime) score was calculated based on participants' hit rates and false alarm rates, following the A scores outlined by Zhang and Mueller (2005). A higher A' score is indicative of better working memory performance, where participants have fewer misses and incorrect key presses.

---

[4]A Principal Component Analysis (PCA) was run on all English proficiency measures. Two components with eigenvalues > 1 emerged, with self-assessed proficiency loading as one component and objective proficiency as a second component. No differences were found between groups on either self-assessed, $F(1, 45) = 2.06$, $p = .16$, $\eta^2 = .04$, or objective proficiency, $F(1, 45) = 2.73$, $p = .11$, $\eta^2 = .06$, but English proficiency scores were positively correlated, $r = .46$, $p < .001$.

Socio-economic status

A composite SES score was captured from the CLiP-Q and calculated from an index of 12 household assets (0 = no, 0.5 = yes, summed), annual household income (value range according to tax brackets; 1 = less than R195,850 to 6 = R708,311 and above), and maternal and paternal level of education (1 = primary school to 6 = Master's or higher). Participants' type of primary and secondary school was used as an additional indicator of SES (1 = government township/rural school, 2 = government suburban school, 3 = private school, summed across primary and secondary schooling; range 2–6).[5]

Procedure

Students who volunteered to participate completed the CLiP-Q online prior to attending the experimental session. Experimental sessions were carried out at a university with a researcher and one participant at a time. Each session lasted between one-to-two hours and took place in one sitting with regular breaks. The verbal fluency task was audio recorded and subsequently transcribed, the 2-back, ANT, and AMDT tasks were completed on PsychoPy (Peirce, 2007), and the WAIS-IV^{SA} vocabulary subtest and ART-2 were manually recorded using scoring templates from the assessment manuals. The order of assessments was counterbalanced across one of four conditions where each participant completed the next consecutive order. To ensure uniformity, however, all participants began the session with the verbal fluency task and the fourth task was always the vocabulary subtest. Participants gave their informed consent prior to completing the CLiP-Q online and again before testing, and received monetary compensation.

Data analyses

Initial analyses to assess normality and control for confounds were carried out. Using ANOVAs, we compared the continuous scores of each background measure across language group (see Table 2). No differences were found in any of the self-assessed English proficiency measures, the objective English proficiency measures, and working memory (all ps > .05), suggesting that the groups were comparable in these abilities. However, both SES measures presented differences between groups, where the L1-English group had a higher SES score than the L1-Zulu group. Unfortunately, these SES differences are reflective of the inherent inequality between the two sample groups that were selected from South Africa (Cornwell & Inder, 2008; Leibbrandt, Finn & Woolard, 2012), and therefore, SES cannot be teased apart from the statistical analyses.[6]

For the RL analyses, only accurate data was used. Furthermore, outliers were identified using the *boxplot* function, where data points are considered outliers if they lie outside 1.5 times the interquartile range above the upper quartile and below the lower quartile (Becker, Chambers & Wilks, 1988). A total of 92 (5.86%) outliers were removed from the pure target word RL index (monitoring process), and 134 (8.54%) outliers were removed from the pure sentence RL index (revision process). These two pure RL indexes were analysed through linear mixed effects models (LME) with the *lmer* function of the *lme4* R package (Bates et al., 2021), whereas accuracy, as it is a binomial measure (0 or 1), was analysed through mixed-effects logistic regression (MELR) models with the *glmer* function of the same package. All models included Participants and Items as random factors, and Condition (expected vs. unexpected for monitoring; expected-congruent vs. expected-incongruent vs. unexpected-congruent vs. unexpected-incongruent for revision), Group (L1-English vs. L1-Zulu), and the centered values of IC, as the fixed factors. Thus, the full fixed structure of each model was always a three-way interaction (condition*group*IC) as well as the combination of all its lower-level interactions and main effects. Using the *simr* package (Green & MacLeod, 2016), we ran power analyses with 1,000 simulations and alpha set at .05 to assess whether the sample size was sufficient to detect effects in our models.

To determine the optimal structure for the random and fixed components, we followed the procedure outlined by Zuur, Ieno, Walker, Saveliev and Smith (2009). We first looked for the best random structure using maximum likelihood, while the full fixed structure (i.e., three-way interaction) was retained (Barr, Levy, Scheepers & Tily, 2013). More specifically, the random structure was tested by running an ANOVA between all possible models containing the various combinations of intercepts and/or slopes, using group and condition as random slopes, while keeping the full fixed structure. The converging model that contained the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) values was selected. Once the best random structure was identified, we then tried to obtain the best fixed structure. To do this, we ran stepwise model comparisons from the most complex model (i.e., three-way interaction) to the simplest model (i.e., main effects), by selecting the significant χ^2 test for the log-likelihood, using maximum likelihood. Third, F and p values were provided by the ANOVA function of the *lmerTest* package using Satterthwaite's method and χ^2 and p values were provided by the ANOVA function of the *car* package using Type II Wald chi-square tests (Fox et al., 2021). To qualify the two-way interactions, the testInteractions function of the *phia* package (De Rosario-Martínez, 2015) and emmeans function of the *emmeans* package (Lenth, Buerkner, Herve, Love, Riebl and Singmann, 2021) were used for post-hoc analyses with Bonferroni correction.

Results

Our results are divided into two sections, depending on the AMDT structure. First, we present the RL results obtained after the presentation of the critical sentence containing the target word (i.e., pure target word RL index), regarding the comprehension monitoring process. Subsequently, we present the two measures assessed in the comprehension sentence (i.e., pure sentence RL index and accuracy), in relation to the revision process. In this paper we focused on the fixed effects of the models but summary details regarding model fit and random effects of each model are provided in Appendix B, as well as Pearson correlations between each variable in Appendix C. Results of the power analyses confirmed that the sample size for each model had sufficient power: pure target word RL index = 100% (CI = 99.63–100%), pure sentence RL index = 88.90% (CI = 86.79–90.78%), and accuracy = 100% (CI = 99.63, 100%) (Brysbaert, 2019).

---

[5]A PCA was also applied to the SES measures, where it was found that all indicators loaded as one component (eigenvalue > 1). L1-English participants had a significantly higher SES index score than L1-Zulu participants, $F(1, 45) = 9.11$, $p < .01$, $\eta^2 = .17$.

[6]When SES was added to the statistical models, the overall results remained the same.

Target word in the text: Monitoring process

An LME model with Participants and Items as the random factors, and Condition (expected vs. unexpected), Group (L1-English vs. L1-Zulu), and IC as the fixed factors, was run on the pure target word RL index. A significant main effect of condition emerged, $F(1, 1437.76) = 83.26$, $p < .001$, $\eta_p^2 = .05$, which showed that participants were faster to respond to expected words ($M = 1.17$, $SE = .03$) compared to unexpected words ($M = 1.43$, $SE = .03$). A marginal effect of group was also found, $F(1, 46.87) = 3.54$, $p = .07$, $\eta_p^2 = .07$, where L1-Zulu participants ($M = 1.35$, $SE = .04$) were marginally slower to respond to the critical sentences than L1-English participants ($M = 1.25$, $SE = .04$). More importantly, a two-way interaction of group × condition emerged, $F(1, 1437.76) = 9.32$, $p < .01$, $\eta_p^2 = .01$ (see Figure 1), where post-hoc comparisons revealed a significant difference between groups for the unexpected, $\chi^2(1) = 9.73$, $p < .01$, but not the expected condition, $\chi^2(1) = 0.01$, $p \approx 1.00$. Specifically, while both groups were equally efficient at responding to words in the expected condition, L1-Zulu participants were slower to respond to the critical sentence than L1-English participants when this sentence contained an unexpected word. No other effects were significant (all $ps > .56$). These findings suggest that while both groups were equally efficient at comprehending coherent information, the L1-Zulu group took longer to monitor inconsistent information in English than the L1-English group, suggesting a less efficient monitoring process. In addition, none of the monitoring process effects were qualified by individual differences in IC.

Comprehension sentence: Revision process

Pure sentence RL index

A second LME model with Participants and Items as the random factors, and Condition (expected-congruent vs. expected-incongruent vs. unexpected-congruent vs. unexpected-incongruent for revision), Group (L1-English vs. L1-Zulu), and IC as the fixed factors, was performed on the pure sentence RL index. The main effect of condition was significant, $F(3, 1394.88) = 12.04$, $p < .001$, $\eta_p^2 = .03$, where post-hoc comparisons showed that the pure sentence RL index was higher on the unexpected-incongruent condition ($M = 1.54$, $SE = .06$) in comparison to the other three conditions: (a) expected-congruent ($M = 1.23$, $SE = .06$), $t(1405) = -5.30$, $p < .001$, (b) expected-incongruent ($M = 1.22$, $SE = .06$), $t(1410) = -5.27$, $p < .001$, and (c) unexpected-congruent ($M = 1.33$,

$SE = .06$), $t(1410) = -3.66$, $p < .01$. No other differences were found between these conditions ($ps > .31$). A main effect of group also emerged, $F(1, 46.11) = 13.85$, $p < .001$, $\eta_p^2 = .23$, where L1-English participants ($M = 1.22$, $SE = .06$) were significantly faster to respond to the comprehension sentences than L1-Zulu participants ($M = 1.43$, $SE = .06$). A third main effect of IC emerged, $F(1, 45.67) = 9.48$, $p < .01$, $\eta_p^2 = .17$, where participants with higher conflict costs (i.e., lower IC) were faster to respond to the comprehension sentences in general. In addition, the two-way interaction of group × condition was marginally significant, $F(3, 1395.88) = 2.46$, $p = .06$, $\eta_p^2 = .01$, and the interaction of condition × IC was significant, $F(3, 1390.83) = 2.84$, $p < .05$, $\eta_p^2 = .01$. No other interactions were significant (all $ps > .17$).

Post-hoc comparisons in the interaction of group and condition revealed that compared to the L1-English group, the L1-Zulu group took longer to respond to the comprehension sentences when presented with incongruent information: expected-incongruent, $t(185) = -3.15$, $p < .05$, and unexpected-incongruent, $t(230) = -3.72$, $p < .01$, conditions; but this did not happen when information was congruent: expected-congruent, $t(171) = -1.01$, $p = .97$, and unexpected-congruent, $t(187) = -2.40$, $p = .25$, conditions, where L1-Zulu participants responded more similarly to the L1-English group (see Figure 2). Therefore, as it occurred in the monitoring process, our results indicate that the two groups did not differ when the story was completely coherent (i.e., expected-congruent), and they had to confirm the information in the comprehension sentence (i.e., unexpected-congruent). In contrast, L1-Zulu participants were less efficient than L1-English participants when the comprehension sentence required the disconfirmation of the presented information (i.e., expected-incongruent and unexpected-incongruent) as well as when it demanded revision of the initial interpretation followed by content that did not align with the unexpected information (i.e., unexpected-incongruent).

Moreover, pairwise comparisons in the interaction of condition and inhibitory control showed significant differences between the regression slopes of the four conditions, $\chi^2(3) = 8.52$, $p < .05$, where participants with higher inhibitory cost (lower IC) took longer in the unexpected-incongruent condition compared to the other three conditions. In contrast, participants with lower inhibitory cost (higher IC) showed no differences between conditions (see Figure 3).

Accuracy

An MELR model with Participants and Items as the random factors, and Condition (expected-congruent vs. expected-incongruent vs. unexpected-congruent vs. unexpected-incongruent for revision), Group (L1-English vs. L1-Zulu), and IC as the fixed factors, was run on accuracy to the comprehension sentence. Similar to the previous measure, a significant main effect of condition emerged, $F(3, 66.47) = 22.16$, $p < .001$, where post-hoc comparisons showed lower accuracy on the unexpected-incongruent condition, ($M = .72$, $SE = .11$), compared to the other three conditions: (a) expected-congruent, ($M = .91$, $SE = .16$), $z = 7.10$, $p < .001$, (b) expected-incongruent, ($M = .86$, $SE = .14$), $z = 5.22$, $p < .001$, and (c) unexpected-congruent, ($M = .87$, $SE = .14$), $z = 5.53$, $p < .001$. There was also a significant main effect of group, $F(1, 5.78) = 5.78$, $p < .01$, with higher accuracy scores for L1-English participants ($M = .86$, $SE = .09$) compared to L1-Zulu participants ($M = .81$, $SE = .09$). No other conditions or interactions were significant ($ps > .19$). Therefore, all participants performed worst in the most difficult revision condition (i.e.,
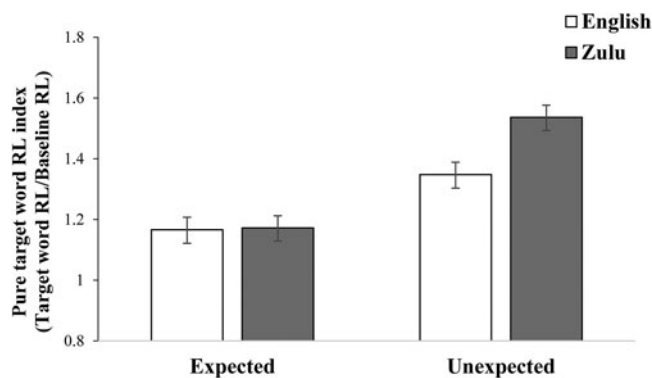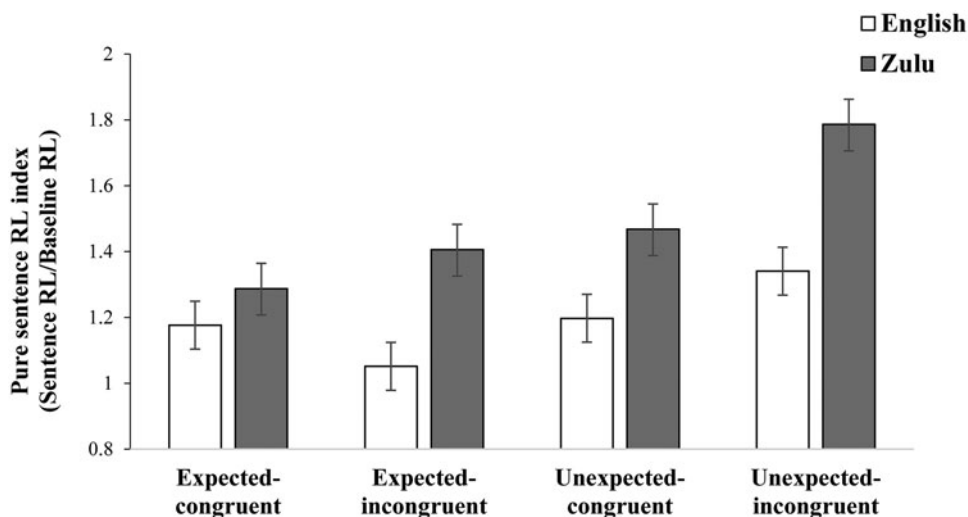

Fig. 1. Mean and standard error values of the pure target word response latency index obtained after listening to the critical sentence of the AMDT, between the language groups across each of the expectancy conditions.

Fig. 2. Mean and standard error values of the pure sentence response latency index obtained after listening to the comprehension sentence of the AMDT, between the language groups across each of the four expectancy and congruency conditions.
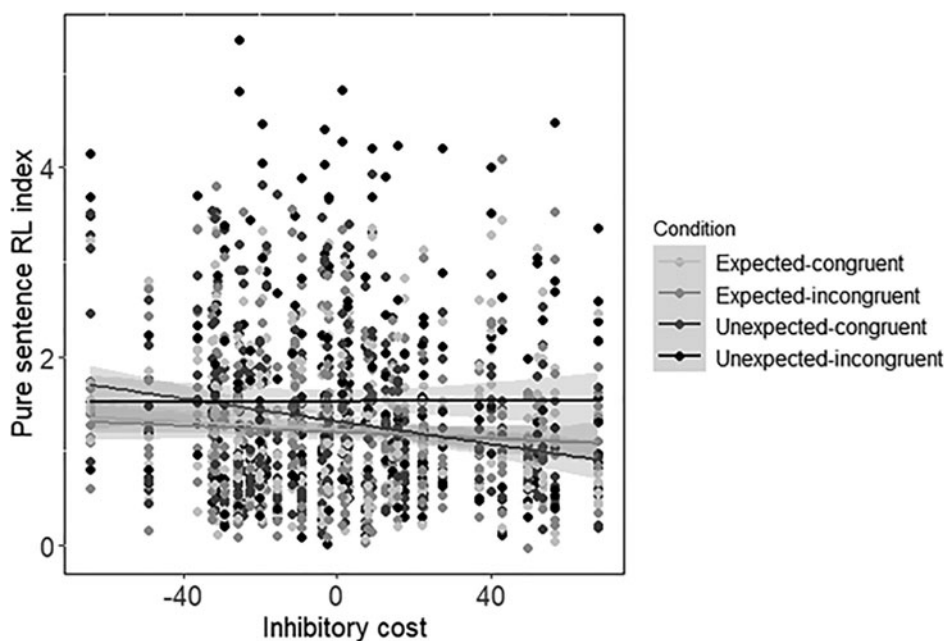

Fig. 3. Scatter linear graph with mean comprehension sentence response latency scores, representing the interaction between condition and inhibitory control.

unexpected-incongruent) and the L1-English group manifested better comprehension (and therefore, better revision) than the L1-Zulu group.

Discussion

This study is the first to evaluate high-level listening comprehension in young South African adults with advanced proficiency in English – the language of tertiary education and the country's *lingua franca* – but who have differing L1s that either match (English) or differ from (Zulu) the language of the listening comprehension task. Our first goal was to understand how these groups monitor and revise inferential information while listening to narrative stories. We were additionally interested in exploring whether individual differences in IC explained any effects associated with listening comprehension processes.

Monitoring in high-level listening comprehension

The comprehension monitoring process was evaluated by manipulating the expectancy of the target word presented in the critical sentence, which either aligned (expected) or misaligned (unexpected) but was still plausible with the generated inference made from the establishing context. Accordingly, the pure target word RL index for the critical sentence was used to evaluate whether comprehenders had generated the knowledge-based inference facilitated by the establishing context, and subsequently detected a mismatch between this prediction and the unexpected word.

Previous studies on comprehension monitoring within both the reading modality and auditory-verbal and visual-pictorial modality at the sentence and discourse level have demonstrated a general expectancy effect (Baker, 1989; Hu & Jiang, 2011; Pérez et al., 2016, 2019, 2020; Vorstius, Radach, Mayer & Lonigan, 2013). In line with these findings, increased RLs in the unexpected compared to the expected condition confirmed that our participants had generated an initial inference and were able to monitor their comprehension across the auditory modality (confirming Hypothesis 1). This suggests that the initial inference generated from the establishing sentence was more likely to be activated in the situation model than the alternative interpretation when participants encountered the target word in the critical sentence. Notably, these results also indicate that a similar comprehension monitoring process is evident across various forms of content presentation, whereby the ability to monitor inferential information across different modalities is efficient and largely modality independent.

Importantly, we also observed that L1 background played a role in comprehension monitoring efficiency when there was prediction interference. As predicted (Hypothesis 2), L1-Zulu participants were slower to process sentences containing the unexpected word compared to L1-English participants, while there were no group differences in the expected condition. Thus, monitoring was less efficient for comprehenders who, although highly proficient, had a different L1 to that of the task when there was misalignment between the target word in relation to the previously presented content. A similar reading time effect was observed in an English monolingual sample with a slightly different paradigm involving three conditions evaluating comprehension monitoring (a neutral condition, which contained general information following the establishing content; a no-revise condition, which was similar to the expected condition containing a likely concept; and a revise condition, which contained an unexpected but plausible concept) (Pérez, Cain, Castellanos & Bajo, 2015). Similar results have also been found in behavioural and electrophysiological studies evaluating comprehension monitoring, where highly proficient L2 speakers are less sensitive to upcoming content and do not anticipate information to the same extent compared to L1 speakers, likely because more time is required to deal with unexpected information in a non-foundational language (Chambers & Cooke, 2009; Kaan, Kirkham & Wijnen, 2016; Martin, Thierry, Kuipers, Boutonnet, Foucart and Costa, 2013). This effect may be related to typological differences between languages, though (Foucart, Martin, Moreno & Costa, 2014).

Comparisons can further be drawn with acoustic studies assessing monosyllabic word comprehension in the presence of noise and/or reverberation interference. Monolinguals consistently outperform highly proficient L2-English bilinguals during English speech perception in noise tasks, suggesting that meaning extraction is compromised during speech signal detection in L2 listening (Rogers, Lister, Febo, Besing & Abrams, 2006; Shi, 2010; Tabri, Chacra & Pring, 2011). Importantly, these results are not explained by proficiency but rather by participants' L1 background as it relates to the language of the task. Comprehension monitoring may therefore draw on automatised processing which may be harder to attain if content is not presented in the L1.

Earlier theories of sentence processing postulated that non-L1 comprehenders have a Reduced Ability to Generate Expectations (RAGE; Kaan, Dallas & Wijnen, 2010; Grüter,

Rohde & Shafer, 2014) and recent research suggests that the mechanisms underlying sentence processing, and prediction in particular, is comparable in L1 and L2 comprehenders, where results should be explained in terms of individual differences between groups (Kaan, 2014). Accordingly, there does appear to be a processing advantage for individuals with an L1 that matches the language of the task, even when all participants have advanced proficiency, which is attributable to a heightened processing sensitivity to variation within the semantic context. Interestingly, 12 (54.5%) L1-Zulu participants reported English as being their most proficient language. This result may too be a consequence of the quantity and quality of English input and usage between the groups, given that L1-English participants reported to use English with a variety of interlocutors (e.g., family, friends) more so than L1-Zulu participants (see Appendix A). Further research exploring questions of typological distance between languages as well as input and usage quantity and quality in multilingual contexts as it relates to the monitoring process in listening comprehension should therefore be avenues of exploration.

Revision in high-level listening comprehension

During successful comprehension of unfolding narratives, the process of revision is required such that comprehenders need to review the existing text interpretation and update their situation model so that the new and potentially unexpected content can become meaningfully integrated into the discourse context. This process is cognitively demanding and largely dependent on the proficiency and language experience of comprehenders (Kendeou et al., 2013; Pérez et al., 2015, 2016, 2020). Until now, however, no study has investigated whether L1, as either corresponding with (English) or differing from (Zulu) the language of the task, has an effect on the revision process under the condition that English proficiency for the whole sample is advanced. Revision was assessed by combining the expectancy of the target word presented in the critical sentence (expected vs. unexpected) with the congruency of the critical information presented in the comprehension sentence (congruent vs. incongruent), by means of both the pure sentence RL index in and accuracy to the comprehension sentence. Crucially, revision was only required after the presentation of the unexpected target word, to (dis)confirm the comprehension sentence.

In line with the general predictions, we found that RLs were higher (Hypothesis 3) and accuracy lower (Hypothesis 4) on the unexpected-incongruent condition ("*turtles*" → "*feathers*") in comparison to the other three conditions, with no differences between them. These results indicate that participants were least efficient when they were required to disconfirm information that was coherent with their initial inference (e.g., "*feathers*" → "*ducks*"). In fact, this was considered the most difficult revision condition due to the need to be able to discard no longer probable information. In contrast, participants were able to quickly and accurately respond to information that was either coherent with their initial interpretation or incoherent with this interpretation but confirmed by the comprehension sentence.

Using the same paradigm but assessing reading comprehension, Pérez et al. (2016) found that readers spent less time reading congruent compared to incongruent sentences but only for narratives preceded by an expected (but not unexpected) concept. Furthermore, when text containing an unexpected concept was followed by a congruent (but not incongruent) sentence, readers

were slower to respond than when it contained an expected concept. Crucially, for listening comprehension, the only difference found relates to the condition where incongruent content follows an unexpected concept. While for reading comprehension, it appears that prior context facilitates (or impedes) reading when narrative consistency is maintained (or disrupted), for listening comprehension, the effect is evinced only in the most difficult revision condition involving an unexpected target concept followed by an incongruent comprehension sentence. This supports the claim that revising content that is disconfirmed by one's initial inference followed by a further iteration of incongruency is a cognitively taxing and difficult comprehension process.

Importantly, some research has found a prolonged influence of initially encoded information from early parts of a narrative text, such that the initial content exerts an effect on comprehension even when newer content is qualified as the narrative unfolds (Guéraud et al., 2005; Raap & Kendeou 2007), while other research suggests that comprehenders can successfully amend their knowledge to reflect new states of affairs that emerge over the course of a narrative text (de Vega, 1995; Diakidoy, Kendeou & Ioannides, 2003). Depending on the type and difficulty of conditions presented, both arguments hold. Comprehenders are sufficiently able to modify earlier representations of a text with updated content, specifically when there is only one disruption to coherence, but experience processing difficulties when listening to narratives that include prediction interference across two instances of coherence.

In terms of language differences, L1-English comprehenders were significantly faster than L1-Zulu comprehenders to respond to the comprehension sentence. Importantly, a two-way interaction of group × condition also emerged for RLs, where the L1-Zulu group took longer to respond to the comprehension sentence than the L1-English group in the expected-incongruent ("ducks" → "shells"), and unexpected-incongruent ("turtles" → "feathers") conditions, compared to the easiest expected-congruent condition ("ducks" → "feathers", almost confirming Hypothesis 5). The same was also predicted for the unexpected-congruent condition ("turtles" → "shells"), but although means suggested a similar RL trend, this effect was not significant ($p = .25$). Overall, our findings indicate that non-L1 comprehenders experienced more difficulties than L1 speakers in dealing with interference prediction, especially when the hardest revision condition was involved during high-level listening comprehension. These results are consistent with Pérez et al. (2019), who found that participants were less efficient at revising their situation model during a reading comprehension task when it was presented in their L2 compared to their L1, even though these comprehenders were highly proficient in both languages.

For the accuracy analysis, general scores were relatively high, with L1-English and L1-Zulu participants achieving accuracy scores of 85.8% and 81.6%, respectively. However, L1-English comprehenders were more accurate than L1-Zulu comprehenders, suggesting a more accurate inferential comprehension once more in favour of L1-comprehenders (partially confirming Hypothesis 6). Nonetheless, no interaction with condition emerged, so these differences did not seem to be due to difficulties in the revision process.

A possible way to address these differences may be accounted for in terms of metacognition – an awareness and evaluation of one's own cognitive processes and thoughts (Flavell, 1979; Fleming & Lau, 2014). Indeed, metacognition has been found to be lower (Folke, Ouzia, Bright, De Martino & Filippi, 2016)

or equivalent (Filippi, Ceccolini, Periche-Tomas & Bright 2020) in bilinguals compared to monolinguals. A preliminary interpretation of these findings proposes that self-evaluation of performance may be related to the language background of participants, such that confidence is reduced in relation to the number of participants' spoken languages. Applied to this study, it is likely that L1-Zulu participants were less confident in their judgement of auditorily presented information tapping revision purely because English is not their L1 and/or they are regularly exposed to numerous languages, making them more sensitive to resolving conflict and accepting inaccurate information. While the L1-Zulu participants have the linguistic and cognitive resources to deal with prediction interference, and indeed perform remarkably well on other language tasks and reading comprehension (see Table 2), they may be less assured in their interpretation of auditory content, specifically when it is complex and requires the cancellation or questioning of previously generated knowledge and/or inferences. While metacognition was not directly assessed in this study, it is worth exploring its influence in future research to shed light on the effect of self-awareness on non-L1 comprehension processing, given that our interpretation is speculative.

Inhibitory control in high-level listening comprehension

As a second goal, we explored whether high-level comprehension processes were modulated by individual differences in IC. Sentence processing as explained by language differences may be accounted for by processing loads during L2 comprehension, such that when information is processed in a non-L1 language, it requires additional cognitive resources and may be processed at a lower, usually lexical level, leaving fewer resources for conceptual/semantic processing (Frey, 2005; Perfetti, Yang & Schmalhofer, 2008). Comparably, the process of revising outdated inferential content has been found to involve IC, whereby irrelevant information is efficiently disregarded while pertinent information is at the forefront of one's focus (Pérez et al., 2016). This may be borne out during high-level listening comprehension where great inhibitory capacity is required to efficiently process ongoing content, especially for non-L1 comprehenders. As such, IC was evaluated as a factor modulating listening comprehension performance for L1-English and L1-Zulu groups.

As expected, individual differences of IC were not associated with the comprehension monitoring process (Hypothesis 7). This lack of effect is consistent with previous studies showing that text monitoring of knowledge-based predictive inferences is a more passive process and not very cognitively demanding, especially in highly proficient individuals (Pérez et al., 2015, 2016). In contrast, IC was related to the revision process, suggesting that it is more cognitively demanding than the monitoring process (Pérez et al., 2016, 2020). Accordingly, participants with lower IC were faster at responding to comprehension sentences in general, but no differences in accuracy emerged. While not an expected finding, it may be the case that participants with greater IC were more deliberate and considered when responding to comprehension sentences, thereby being slower while still manifesting an accurate interpretation. However, further research would be necessary to clarify this general effect.

More importantly, IC also interacted with condition, where participants with lower IC (i.e., high inhibitory cost) took longer to respond to the most difficult condition (unexpected-incongruent) when they were required to revise their situation model and replace their initial interpretation for a new one,

compared to the rest of the conditions. In contrast, this was not the case for participants with higher IC, who took a similar amount of time in all conditions (partially confirming Hypothesis 8). These findings suggest that participants with lower IC experience difficulties revising their situation model, whereas higher IC participants perform comparably across conditions. Similarly, Pérez et al. (2020) found that L1-English comprehenders with higher IC performed more accurately on multimodal comprehension tapping the revision process, thereby suggesting that participants with higher IC are better able to revise their situation model when confronted with conflicting information across pictorial and verbal modalities.

Finally, we did not observe an effect of IC as a consequence of L1-groups (disconfirming Hypothesis 9), suggesting that L1-Zulu participants across the IC spectrum were comparable in processing efficiency and accuracy to L1-English participants when answering comprehension sentences. There does not appear to be a relationship between IC as it relates specifically to L1 background in highly proficient young adults. Rather, we found a general effect of IC across all participants, irrespective of their L1 background. A possible explanation may be linked to participants' extensive quantity and high quality of English exposure and use spanning a number of communicative contexts, coupled with the fact that all participants could be classified as multilingual to some degree and might therefore have enhanced executive control from regular language competition (Bialystok, Craik, Green & Gollan, 2009). Notably, none of the participants were characterised as 'typically' monolingual in the sense that they spoke and had experience with only one language (see Appendix A). For instance, of the 10 participants who self-reported as monolingual, they indicated that they spoke between two to five languages ($M = 3.40$, $SD = .84$) and could write in one to three of these ($M = 2.30$, $SD = .67$), suggesting that this group actively uses more than one language. It has also been well established that an individual's languages are concurrently active irrespective of whether only one is being used at any given time (Kroll, Bobb, Misra & Guo, 2008; Wu & Thierry, 2010). Accordingly, suppression of non-target language(s) is a regular cognitive activity for bi/multilinguals, which is thought to enhance executive function beyond the domain of language (Abutalebi & Green, 2007; Stocco, Yamasaki, Natalenko & Prat, 2014). Such cognitive control may be heightened in individuals living in highly linguistically diverse contexts such as South Africa, explaining why we did not observe group differences in IC. In future studies, it would be beneficial to evaluate the role of other (continuous) language factors, such as dominance, English age-of-acquisition, degree of exposure to languages within the ambient context and so forth, since some of these factors have been found to be determinants of executive control and/or comprehension ability in multilingual South Africans (Espi-Sanchis & Cockcroft, 2021; Cockcroft, Wigdorowitz & Liversagem, 2019; Palane & Howie, 2020).

Conclusion

Our results confirm an effect of L1 on listening comprehension, where L1-English participants were more efficient and accurate at performing high-level comprehension processes than L1-Zulu participants. Most strikingly is that the participants' L1 appears to supersede their advanced English proficiency on highly complex listening comprehension involving revision. Having an L1 that is the same as the task being assessed is facilitative to comprehension performance, even when exposure to the task language is

substantial and proficiency is advanced. Furthermore, while there are processing costs associated with comprehension in the auditory modality of linguistically and cognitively demanding information, these do not appear to correspond to costs in other linguistic tasks or reading comprehension, indicating the remarkable ability of learners with advanced English proficiency but who have a different L1. Together, these results reinforce the idea that comprehenders with an L1 similar or dissimilar to that of the task language, but who have high proficiency in the language of the task were faster at revising their situation model when their initial interpretation had changed, because they were better able to suppress the interference generated by their initial prediction equally well. Since this study was conducted on South African university students, the results may only be generalisable to similar linguistically diverse and highly educated populations. Notwithstanding, given that proficiency is considered a primary predictor of comprehension and educational outcomes more generally, this study further illuminates the picture by illustrating that L1 is also important to consider in comprehension performance, especially in linguistically diverse contexts.

References

Abutalebi J and Green D (2007) Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* **20**, 242–275. http://dx.doi.org/10.1016/j.jneuroling.2006.10.003

Arrington N, Kulesz PA, Francis DJ, Fletcher JM and Barnes MA (2014) The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading* **18**, 325–346. doi:10.1080/10888438.2014.902461

Aryadoust V (2019) An integrated cognitive theory of comprehension. *International Journal of Listening* **33**, 71–100. doi:10.1080/10904018.2017.1397519

Babayiğit S and Shapiro L (2020) Component skills that underpin listening comprehension and reading comprehension in learners with English as first and additional language. *Journal of Research in Reading* **43**, 78–97. doi:10.1111/1467-9817.12291

Baker L (1989) Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review* **1**, 3–38.

Barr DJ, Levy R, Scheepers C and Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**, 255–278.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B et al. (2021) *Package "lme4"*. R package version 1.1-27.1. R Core Team. Retrieved from https://cran.r-project.org/web/packages/lme4/lme4.pdf

Becker RA, Chambers JM, and Wilks AR (1988) *The New S Language*. Wadsworth, & Brooks/Cole.

Bialystok E, Craik FIM, Green DW and Gollan TH (2009) Bilingual minds. *Psychological Science in the Public Interest* **10**, 89–129. https://doi.org/10.1177/1529100610387084

Blumenfeld HK and Marian V (2011) Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257. doi:10.1016/j.cognition.2010.10.012

Bodie GD, Worthington DL, Imhof M and Cooper L (2008) What would a unified field of listening look like? A proposal linking past perspectives and future endeavors. *International Journal of Listening* 22, 103–122. doi:10.1080/10904010802174867

Brooks P, Everatt J and Fidler R (2016) *Adult Reading Test 2*. Pearson.

Brysbaert M 2019 How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition* 2, 1–38. https://doi.org/10.5334/joc.72

Bubic A, von Cramon DY and Schubotz RI (2010) Prediction, cognition and the brain. *Frontiers in Human Neuroscience* 4, 1–15. doi:10.3389/fnhum.2010.00025

Carretti B, Borella E, Cornoldi C and De Beni R (2009) Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences* 19, 246–251. doi:10.1016/j.lindif.2008.10.002

Chambers CG and Cooke H (2009) Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory and Cognition* 35, 1029–1040. https://doi.org/10.1037/a0015901

Cockcroft K, Wigdorowitz M and Liversage L (2019) A multilingual advantage in the components of working memory. *Bilingualism: Language and Cognition* 22, 15–29. https://doi.org/10.1017/S1366728917000475

Constitution of the Republic of South Africa (1996) *The Constitution of the Republic of South Africa*. Retrieved from https://www.refworld.org/docid/3ae6b5de4.html

Cook AE and O'Brien EJ (2017) Fundamentals of inferencing during reading. *Language and Linguistics Compass* 11, e12246. doi:10.1111/lnc3.12246

Cornwell K and Inder B (2008) Language and labour markets in South Africa. *Journal of African Economies* 17, 490–525. https://doi-org.ezp.lib.cam.ac.uk/10.1093/jae/ejm037

De Rosario-Martínez H (2015) *Package "phia". R package version 0.2-1*. R Core Team. Retrieved from https://cran.r-project.org/web/packages/phia/phia.pdf

de Vega M (1995) Backward updating of mental models during continuous reading of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 373–385. https://doi.org/10.1037/0278-7393.21.2.373

de Vega M, Urrutia M and Riffo B (2007) Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory, & Cognition* 35, 1410–1421. https://doi-org.ezp.lib.cam.ac.uk/10.3758/BF03193611

Department of Education (DoE) (1997) Language in Education Policy, *Government Gazette* No. 18546. 19 December. 5–7.

Diakidoy IN, Kendeou P and Ioannides C (2003) Reading about energy: The effects of text structure in science learning and conceptual change. Contemporary *Educational Psychology* 28, 335–356.

Dijkstra T, Van Hell J and Brenders P (2015) Sentence context effects in bilingual word recognition: Cognate status, sentence language, and semantic constraint. *Bilingualism: Language and Cognition* 18, 597–613. https://doi.org/10.1017/S1366728914000388

Drummond A (2016) An analysis of language policy versus practice in two South African universities. *Southern African Linguistics and Applied Language Studies* 34, 71–79. https://doi.org/10.2989/16073614.2016.1159522

Espi-Sanchis G, & Cockcroft K (2021) Working memory and multilingualism: balanced language proficiency predicts verbal working memory. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2021.1997901

Fan J, McCandliss BD, Sommer T, Raz A and Posner MI (2002) Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience* 14, 340–347. https://doi.org/10.1162/089892902317361886

Filippi R, Ceccolini A, Periche-Tomas E and Bright P (2020) Developmental trajectories of metacognitive processing and executive function from childhood to older age. *Quarterly Journal of Experimental Psychology* 73, 1757–1773. https://doi.org/10.1177/1747021820931096

Filippi R, Leech R, Thomas MS, Green DW and Dick F (2012) A bilingual advantage in controlling language interference during sentence comprehension. *Bilingualism: Language and Cognition* 15, 858–872. http://dx.doi.org/10.1017/S1366728911000708.

Flavell JH (1979) Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist* 34, 906–911.

Fleming SM and Lau HC (2014) How to measure metacognition. *Frontiers in Human Neuroscience* 8, 1–9. doi:10.3389/fnhum.2014.00443

Flesch R (1949) *The art of readable writing*. New York: Harper and Brothers.

Folke T, Ouzia J, Bright P, De Martino B and Filippi R (2016) A bilingual disadvantage in metacognitive processing. *Cognition* 150, 119–132. doi:10.1016/j.cognition.2016.02.008

Foucart A, Martin CD, Moreno EM and Costa A (2014) Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40, 1461–1469. https://doi.org/10.1037/a0036756

Fox J, Weisberg S, Price B, Adler D, Bates D, Baud-Bovy G, Bolker, B et al. (2021) *Package "car". R package version 3.0-11*. R Core Team. Retrieved from https://cran.r-project.org/web/packages/car/car.pdf

Frey L (2005, January) The nature of the suppression mechanism in reading: Insights from an L1-L2 comparison. In *Proceedings of the Annual Meeting of the Cognitive Science Society* 27, 714–719.

Friesen DC, Luo L, Luk G and Bialystok E (2015) Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience* 30, 238–250, doi:10.1080/23273798.2014.918630

Graesser AC, Singer M and Trabasso T (1994) Constructing inferences during narrative text comprehension. *Psychological Review* 101, 371–395. doi:10.1037/0033-295X.101.3.371

Green P and MacLeod CJ (2016) SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution* 7, 493–498. https://doi.org/10.1111/2041-210X.12504

Grüter T, Rohde H and Shafer AJ (2014) The role of discourse-level expectations in non-native speakers' referential choices. In Orman W and Valleau MJ (eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development*. Cascadilla Press, pp. 179–191.

Guéraud S, Harmon ME and Peracchi KA (2005) Updating situation models: The memory-based contribution. *Discourse Processes* 39, 243–263. https://doi.org/10.1207/s15326950dp3902&3_8

Hu G and Jiang N (2011) Semantic integration in listening comprehension in a second language: Evidence from cross-modal priming. In Trofimovich P and Kim M (eds.), *Applying priming methods to L2 learning, teaching and research: insights from psycholinguistics*. Philadelphia: Benjamins, pp. 199–212.

Kaan E, Dallas AC and Wijnen F (2010) Syntactic predictions in second-language sentence processing. In Zwart J-W and de Vries M (eds.), *Structure preserved. Festschrift in the honor of Jan Koster*. John Benjamins, pp. 207–213.

Kaan E and Grüter T (2021) Prediction in second language processing and learning: Advances and directions. In Kaan E and Grüter T (eds.), *Prediction in second language processing and learning*. John Benjamins, pp. 2–24.

Kaan E (2014) Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism* 4, 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan E, Kirkham J and Wijnen F (2016) Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition* 19, 1–18. https://doi.org/10.1017/S1366728914000844

Kendeou P, Smith ER and O'Brien EJ (2013) Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 854–865. https://doi.org/10.1037/a0029468

Kim Y-SG (2015) Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology* 141, 101–120. http://dx.doi.org/10.1016/j.jecp.2015.08.003

Kincaid JP, Fishburne RP, Rogers RL and Chissom BS (1975) *Derivation of new readability formulas (automated readability index, fog count, and Flesch reading ease formula) for Navy enlisted personnel*. [Research Branch Report

8–75]. Chief of Naval Technical Training: Naval Air Station Memphis. Retrieved from https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=1055&context=istlibrary

Kintsch W (1998) Comprehension: A paradigm for cognition. New York, NY: Cambridge University Press.

Kirchner WK (1958) Age differences in short-term retention of rapidly changing information. Journal of Experimental Psychology 55, 352–358.

Kroll JF, Bobb SC, Misra M and Guo T (2008) Language selection in bilingual speech: evidence for inhibitory processes. Acta Psychologica 128, 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001

Leibbrandt MA, Finn A, Woolard I (2012) Describing and Decomposing Post-apartheid Income Inequality in South Africa. Development Southern Africa 29, 19–34. https://doi.org/10.1080/0376835X.2012.645639

Lenth RV, Buerkner P, Herve M, Love J, Riebl H and Singmann H (2021) Package "emmeans". R package version 1.7.0. R Core Team. Retrieved from https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Mackworth JF (1959) Paced memorizing in a continuous task. Journal of Experimental Psychology 58, 206–211.

Madadzhe RN (2019) Using African languages at universities in South Africa: The struggle continues. Stellenbosch Papers in Linguistics Plus 58, 205–218. https://doi.org/10.5842/58-0-843

Manfredi M, Cohn N, Andreoli MDA and Boggio PS (2018) Listening beyond seeing: Event-related potentials to audiovisual processing in visual narrative. Brain and Language 185, 1–8. https://doi.org/10.1016/j.bandl.2018.06.008

Martin C, Thierry G, Kuipers J-R, Boutonnet B, Foucart A and Costa A (2013) Bilinguals reading in their second language do not predict upcoming words as native readers do. Journal of Memory and Language 69, 574–588. http://dx.doi.org/10.1016/j.jml.2013.08.001

McNamara DS (ed.) (2007) Reading comprehension strategies: Theory, interventions, and technologies. Mahwah, NJ: Erlbaum.

McNamara DS and Magliano J (2009) Toward a comprehensive model of comprehension. Psychology of Learning and Motivation 51, 297–384. https://doi.org/10.1016/S0079-7421(09)51009-2

Melby-Lervåg M and Lervåg A (2014) Reading comprehension and its underlying components in second language learners: A meta-analysis of studies comparing first- and second-language learners. Psychological Bulletin 140, 409–433. http://doi.org/10.1037/a0033890.

Mesthrie R (ed.) (2002) Language in South Africa. Cambridge: Cambridge University Press.

O'Brien E, Cook A and Guéraud S (2010) Accessibility of outdated information. Journal of Experimental Psychology: Learning, Memory, and Cognition 36(4), 979–991.

O'Brien EJ, Rizzella ML, Albrecht JE and Halleran JG (1998) Updating a situation model: A memory-based text processing view. Journal of Experimental Psychology: Learning, Memory, and Cognition 24, 1200–1210. doi:10.1037//0278-7393.24.5.1200

Palane NM and Howie S (2020) A comparison of higher-order reading comprehension performance for different language of instruction models in South African primary schools. Perspectives in Education 37, 43–57. https://doi.org/10.18820/2519593X/pie.v37i1.4

Peirce JW (2007) PsychoPy—Psychophysics software in Python. Journal of Neuroscience Methods 162, 8–13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017

Pérez A, Cain K, Castellanos MC and Bajo T (2015) Inferential revision in narrative texts: an ERP study. Memory and Cognition 43, 1105–1135. https://doi.org/10.3758/s13421-015-0528-0

Pérez AI, Hansen L and Bajo T (2019) The nature of first and second language processing: The role of cognitive control and L2 proficiency during text-level comprehension. Bilingualism: Language and Cognition 22, 930–948. https://doi.org/10.1017/S1366728918000846

Pérez AI, Joseph HSSL, Bajo T and Nation K (2016) Evaluation and revision of inferential comprehension in narrative texts: An eye movement study. Language, Cognition and Neuroscience 31, 549–566. https://doi.org/10.1080/23273798.2015.1115883

Pérez AI, Paolieri D, Macizo P and Bajo T (2014) The role of working memory in inferential sentence comprehension. Cognitive Processing 15, 405–413. https://doi.org/10.1007/s10339-014-0611-7

Pérez AI, Schmidt E, Kourtzi Z and Tsimpli IM (2020) Multimodal semantic revision during inferential processing: The role of inhibitory control in text and picture comprehension. Neuropsychologia 138, 107313. https://doi.org/10.1016/j.neuropsychologia.2019.107313

Perfetti C, Yang C-L and Schmalhofer F (2008) Comprehension skill and word-to-text integration processes. Applied Cognitive Psychology 22, 303–318. doi:10.1002/acp.1419

Pickering MJ and Gambi C (2018) Predicting while comprehending language: A theory and review. Psychological Bulletin 144, 1002–1044. https://doi.org/10.1037/bul0000158

Pretorius EJ and Spaull N (2016) Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa. Reading and Writing 29, 1449–1471. https://doi.org/10.1007/s11145-016-9645-9

Rapp DN and Kendeou P (2007) Revising what readers know: updating text representations during narrative comprehension. Memory, & Cognition 35, 2019–2032.

Rapp DN and Taylor HA (2004) Interactive dimensions in the construction of mental representations for text. Journal of Experimental Psychology: Learning, Memory, and Cognition 30, 988–1001. https://doi.org/10.1037/0278-7393.30.5.988

Rapp DN, van den Broek P, McMaster KL, Kendeou P and Espin CA (2007) Higher order comprehension processes in struggling readers: A perspective for research and intervention. Scientific Studies of Reading 11, 289–312.

Révész A and Brunfaut T (2013) Text characteristics of task input and difficulty in second language listening comprehension. Studies in Second Language Acquisition 35, 31–65. https://doi.org/10.1017/S0272263112000678

Rinck M and Weber U (2003) Who when where: An experimental test of the event-indexing model. Memory, & Cognition 31, 1284–1292. https://doi.org/10.3758/bf03195811

Rogers CL, Lister JJ, Febo DM, Besing JL and Abrams HB (2006) Effect of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. Applied Psycholinguistics 27, 465–485.

Rost M (2005) L2 listening. In Hinkel E (ed.), Handbook of research on second language teaching and learning. Mahwah, NJ: Erlbaum, pp. 503–528.

Rost M (2016) Teaching and researching listening (3rd ed). London, England: Longman.

Shi LF (2010) Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. Journal of Speech, Language and Hearing Research 53, 821–835.

South African Department of Education (2008) Minimum admission requirements for Higher Certificate, Diploma and Bachelor's Degree programmes requiring a National Senior Certificate. Retrieved from https://mb.usaf.ac.za/wp-content/uploads/2016/11/nsc_he_admission_requirements.pdf

Statistics South Africa (2012) Census 2011: Census in brief. Statistics South Africa, Pretoria. Retrieved from http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf

Stocco A, Yamasaki B, Natalenko R and Prat CS (2014) Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. International Journal of Bilingualism 18, 67–92. http://dx.doi.org/10.1177/1367006912456617.

Tabri D, Chacra KMSA and Pring T (2011) Speech perception in noise by monolingual, bilingual, and trilingual listeners. International Journal of Language Communication Disorders 46, 411–422.

Tarchi C, Ruffini C and Pecini C (2021) The contribution of executive functions when reading multiple texts: A systematic literature review. Frontiers in Psychology 12, 716463. https://doi.org/10.3389/fpsyg.2021.716463

Tiego J, Testa R, Bellgrove MA, Pantelis C and Whittle S (2018) A hierarchical model of inhibitory control. Frontiers in Psychology 9, 1–25. https://doi.org/10.3389/fpsyg.2018.01339

Vandergrift L (2007) Recent developments in second and foreign language listening comprehension research. Language Teaching 40, 191–210.

van Dijk TA and Kintsch W (1983) Strategies of discourse comprehension. New York: Academic Press.

van Wyk A (2014) English-medium education in a multilingual setting: A case in South Africa. *International Review of Applied Linguistics in Language Teaching* **52**, 205–220. https://doi.org/10.1515/iral-2014-0009

Vorstius C, Radach R, Mayer MB and Lonigan CJ (2013) Monitoring local comprehension monitoring in sentence reading. *School Psychology Review* **42**, 191–206. https://doi.org/10.1080/02796015.2013.12087484

Wagoner SA (1983) Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly* **18**, 328–346.

Webb V, Lafon M and Pare P (2010) Bantu languages in education in South Africa: An overview. Ongekho akekho! – the absentee owner. *Language Learning Journal* **38**, 273–292. https://doi.org/10.1080/09571730903208389

Wechsler D (2008) *Wechsler Adult Intelligence Scale – Fourth edition* (WAIS–IV). San Antonio, TX: Pearson.

Wechsler D (2014) *Wechsler Adult Intelligence Scale – Fourth South African edition* (WAIS-IVSA). Johannesburg: JvR Psychometrics (Pty) Ltd.

Wigdorowitz M, Pérez AI and Tsimpli IM (2020) A holistic measure of contextual and individual linguistic diversity. *International Journal of Multilingualism*, 1–19. https://doi.org/10.1080/14790718.2020.1835921

Wigdorowitz M, Pérez AI and Tsimpli IM (2022) Sociolinguistic context matters: Exploring differences in contextual linguistic diversity in South Africa and England. *International Multilingual Research Journal* **16**, 345–364. https://doi.org/10.1080/19313152.2022.2069416

Wu S and Ma Z (2016) Suppression and working memory in auditory comprehension of L2 narratives: Evidence from cross-modal priming. *Journal of Psycholinguistic Research* **45**, 1115–1135. https://doi.org/10.1007/s10936-015-9390-2

Wu YJ and Thierry G (2010) Chinese–English bilinguals reading English hear Chinese. *Journal of Neuroscience* **30**, 7646–7651.

Zhang LM (2017) *Metacognitive and cognitive strategy use in reading comprehension: A structural equation modelling approach*. Singapore: Springer Nature Singapore.

Zhang J and Mueller ST (2005) A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika* **70**, 203–212.

Zuur AF, Ieno EI, Walker NJ, Saveliev AA and Smith GM (2009) *Mixed effects models and extensions in ecology with R*. Springer.

Zwaan RA and Radvansky GA (1998) Situation models in language comprehension and memory. *Psychological Bulletin* **123**, 162–185. https://doi.org/10.1037/00332909.123.2.162