THE EFFECTS OF ACCENT FAMILIARITY ON NARRATIVE RECALL IN NOISE

Andreas Krug, Ghada Khattab and Laurence White

School of Education, Communication and Language Sciences, Newcastle University, UK
a.krug2@newcastle.ac.uk, ghada.khattab@newcastle.ac.uk, laurence.white@newcastle.ac.uk

ABSTRACT

Listeners frequently hear spoken utterances in different accents. It has been suggested that L2 accents are processed and represented in less lexical and semantic detail by L1 listeners. This has been related to listeners' lower expectations of L2 linguistic proficiency, but the impact of other factors, such as accent familiarity, has not been fully explored. Here we test whether the L2 recall disadvantage extends to unfamiliar L1 accents, to which reduced proficiency expectations should not apply. Tyneside (UK) and New Zealand listeners performed a change detection task, during which they heard story pairs in Tyneside and New Zealand English. Stories were masked with multi-speaker babble noise to increase task difficulty. Accent did not mediate New Zealand listeners' recall, but there was a familiarity benefit for Tyneside listeners, whose recall was more accurate for their own accent. Thus, similar to L2 accent studies, L1 accent recall can be mediated by familiarity.

Keywords: speech perception, recall, familiarity, regional accents, noise masking

1. INTRODUCTION

This paper investigates how narrative recall is affected by accent familiarity when stories are presented in noise. Past research has shown that short-term recall of stories is affected by the L1/L2 status of the narrator [1]. Specifically, word changes between an orally narrated story and a corresponding transcript were detected more reliably if the narrator was an L1 rather than L2 speaker of English, though only if participants were asked to listen for comprehension rather than memory [1].

These findings have been interpreted within the framework of "less-detailed" or "good-enough" representations [1, 2, 3]: thus, the processing and subsequent representation of linguistic stimuli is only as detailed as required for the task at hand. This account predicts that factors such as task instructions and linguistic structure will influence the processing and representation of narrations. For example, stored representations during the task will be more detailed if listeners are instructed to listen for memory rather than comprehension. In terms of linguistic structure, focus encourages more detailed representations. Thus, semantically related changes are detected more frequently when presented in focus versus out of focus [4]. Only in linguistic focus are representations detailed enough to detect the changes. Lev-Ari and Keysar's [1] results suggest that speaker identity also influences processing. When stories were narrated by L1 speakers of English, listener performance in a change detection paradigm was better than when the narrator was an L2 speaker, suggesting more detailed lexical and semantic representations of L1 narrations [1]. Importantly, this L1/L2 accent effect was not found if listeners were instructed to focus on potential word changes rather than comprehension beforehand: they presumably engage more detailed processing under such task conditions, regardless of speaker accent [1].

It remains unknown whether the L2 speech recall disadvantage extends to different L1 regional accents. Lev-Ari and Keysar [1] suggest that the poorer recall of L2 speech is due to listeners' expectation that L2 speakers will be less proficient in their language use than L1 speakers. As a result, listeners process and represent L2 speech in less detail to make the processing mechanism more efficient [5, 6]. Less proficient language use would not be expected from speakers of different regional accents of the same L1. However, listeners also hold varying expectations about regional accents in terms of prestige and pleasantness [7, 8]. Moreover, listeners tend to be better at processing familiar rather than unfamiliar accents – as indexed, for example, by lexical access – especially under adverse listening conditions [9, 10, 11]. This disparity could extend to narrative recall if less successful lexical processing results in poorer lexical and semantic representations of speech. Given these considerations, poorer recall of unfamiliar L1 accents is plausible.

A partial replication of Lev-Ari and Keysar's [1] study found no differences in recall due to variation in Spanish L1 accents [12]. However, design differences may have influenced outcomes: for example, listeners had to recall the changed words in Lev-Ari and Keysar [1] while Frances et al. [12] only analysed change detection accuracy.

The current study specifically investigates the effect of accent familiarity on recall. Adverse listening conditions are employed to increase task difficulty and thus potentially elicit effects of accent familiarity that might not otherwise emerge [9]. The study uses a change detection and recall paradigm similar to that of Lev-Ari and Keysar [1] to address the research question: How does familiarity with different L1 accents affect recall of narratives spoken in these accents under adverse listening conditions?

2. METHOD

2.1. Participants

Two groups of participants were recruited via universities and social media platforms: native Tynesiders (UK) and native New Zealanders. A pre-screening questionnaire was used to confirm that the participants had been born and raised in Tyneside or New Zealand [13]. These locations were chosen as providing participant groups unlikely to be familiar with each other's accents. Participants received course credit or vouchers as a token of appreciation. Data from 42 Tynesiders (mean age: 28.6 years) and 37 New Zealanders (mean age: 24.5 years) were analysed. None of the participants reported any diagnosed speech, language or hearing difficulties. Accent familiarity was tested by means of an accent matching task as well as a demographic and language background questionnaire (Section 2.3), which confirmed that the two groups were relatively unfamiliar with each other's accents.

2.2. Stimuli

The stimulus set for the recall task consisted of story pairs differing only in a single target word [see Sturt et al., 4, for a similar design]. Each story consisted of three sentences. For experimental pairs, the target occurred in the second sentence, although never sentence-initially. As the example below shows, the changed target is semantically related to the original one in the 'related' version of the story (*rubies* → *diamonds*). In the 'unrelated' version, the two targets are unrelated (*rubies* → *drugs*). Semantically related targets could be subsumed under one hyperonym (e.g. *gems*). For filler pairs, the target occurred in the first or the third sentence.

Example story: *Sam and Kate made an interesting discovery yesterday. They found a small chest full of* <u>*rubies/diamonds/drugs*</u> *in their attic. They had no idea where it came from or who put it there.*

Stories were recorded at Newcastle University (UK) and Canterbury University (New Zealand) in similar studio conditions. For each accent, there were two middle-aged female speakers, judged by phonetically trained assessors to be representative of their accents. The Tyneside English (TE) speakers had lived in or close to Newcastle upon Tyne (UK) for most of their lives. The New Zealand English (NZE) speakers were of European descent, born and raised in the Christchurch region and had lived in New Zealand most of their lives. European NZE accents are relatively homogeneous [14], which is why those two speakers were deemed representative of not just Christchurch accents, but of European NZE in general.

For each accent, there were 12 experimental and 8 filler story pairs. All stimuli were masked with amplitude-stable multi-speaker babble noise [15], using an adapted Praat script originally developed by McCloy [16]. The amplitude of the stories was first normalised to 65dB at the sentence level. The stories were then overlaid with the babble noise at a signal-to-noise ratio of 0dB. The noise signal was faded in and out for 500 ms before and after the speech signal.

2.3. Procedure

The recall task was part of a larger study, conducted online via LabVanced [17]. The study included: 1. a headphone check; 2. a lexical decision task (reported elsewhere); 3. the recall task; 4. the accent matching task; 5. the demographic and language background questionnaire.

The recall task included 40 story pairs (24 experimental), preceded by four practice pairs. For each speaker, there were 10 trials (6 experimental), 5 with changes and 5 without, as shown in Table 1. The story pairs were blocked by accent, with speaker nested under accent. The order of speakers was varied between participants via a Latin-square design.

Experimental Trials			Filler Trials	
un-changed 2	related change 2	unrelated change 2	un-changed 3	changed 1
Total = 6			Total = 4	
Grand total = 10				

Table 1: Trials per speaker per participant during the recall task

On each trial, participants first saw centrally located fixation crosses for 500 ms. They then heard the first story, followed by another 500 ms of fixation crosses. Then they heard the second story, which: a. was identical to the first one; b. included a related word change (*rubies* → *diamonds*); c. included an unrelated word change (*rubies* → *drugs*). Participants had to respond, via key presses, whether the two stories were identical or different. If they responded "identical", they moved on to the next trial. If they responded "different", they saw the second story on a new screen, including two text boxes for participants to input the changed word from the second story and the original word from the first story.

2.4. Data Analysis

The data were analysed within the 'tidyverse' package in R [18, 19]. The data analysis was conducted in two steps. First, participants' key press accuracy was analysed, i.e., if they correctly responded that the two stories were identical or different. Second, the accuracy of the correction was analysed. This measure was only considered for trials which did include a change and for which the participants had pressed the correct key initially.

Logistic mixed-effects models were used for the statistical analysis of both datasets, which consisted of 1896 and 955 trials for the key press and correction accuracy, respectively. The following predictors were included in the regression models:

- **Accent**: TE *vs* NZE.
- **Story pair**: unchanged *vs* related change *vs* unrelated change. For correction accuracy, unchanged story pairs were irrelevant and this level of story pair was dropped.
- **Participant origin:** Tyneside *vs* New Zealand.
- All interactions (two-way and three-way) between accent, story pair and participant origin were included.

The models further included two control predictors:

- **Block** distinguished the first and second five trials in a specific speaker's voice to check for adaptation effects.
- **Articulation rate**, in syllables per second (z-transformed).

The random effects structure of the models included random intercepts for participant and item, plus random slopes for accent by participant.

For both key press and correction accuracy, the analysis procedure started with the full model, provided below in 'lme4' syntax [20]:

accuracy ~ accent x story pair x participant origin +
block + articulation rate +
(1 + accent|participant) + (1|item)

Significant effects were identified by means of log-likelihood model comparisons, using the 'afex' package in R [21]. Predictors that did not reach significance were dropped. This resulted in the following models, for which Bonferroni-corrected pairwise comparisons were conducted with the 'emmeans' package [22]:

accuracy_key_press ~ accent x story pair x
participant origin +
(1 + accent|participant) + (1|item)

accuracy_correction ~ story pair +
(1| participant) + (1|item)

3. RESULTS AND DISCUSSION

3.1. Key Press Accuracy

Overall, 1388 trials (73.2%) were correct and 508 (26.8%) were incorrect. There was a significant effect of story pair, $\chi^2(2) = 107.0$, $p < .001$, along with interactions between accent and participant origin, $\chi^2(2) = 9.5$, $p = .002$, and between accent, story pair and participant origin, $\chi^2(2) = 6.6$, $p = .038$. Regarding story pair, pairwise comparisons showed significant differences between unchanged and related story pairs, $\beta = 5.8$, $SE = 1.2$, $p < .001$, as well as unchanged and unrelated story pairs, $\beta = 5.5$, $SE = 1.1$, $p < .001$, but not between related and unrelated story pairs, $\beta = 1.0$, $SE = 0.1$, $p > .999$.

Regarding the interaction between accent and participant origin, participants' key press accuracy was higher for their own accent. Pairwise comparisons supported this for the difference between TE and NZE speakers in participants from Tyneside, $\beta = 2.1$, $SE = 0.8$, $p = .039$. The accuracy difference between accents was not significant for New Zealand participants, $\beta = 0.7$, $SE = 0.3$, $p = .418$.

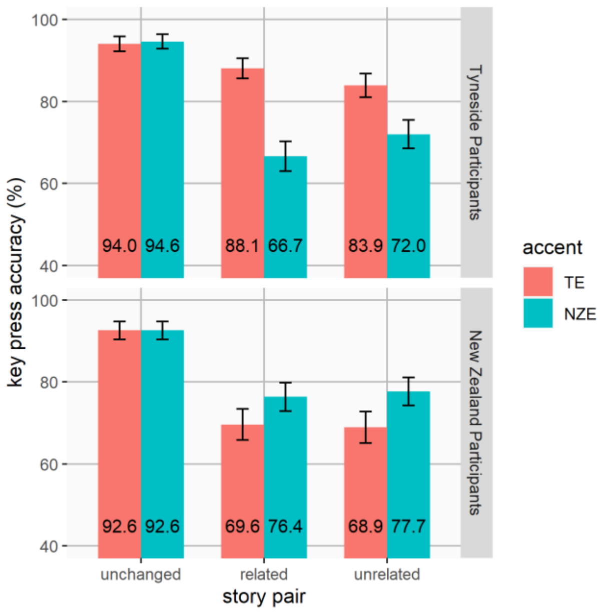

Figure 1: Key press accuracy by participant origin, story pair and accent (error bars: ± one SE)

Figure 1 illustrates the three-way interaction between accent, story pair and participant origin. The main effect of story pair is obvious in that accuracy was highest for unchanged pairs. Pairwise comparisons show that the two-way interaction between accent and participant origin is further modulated by story pair: thus, whilst accuracy did not vary by speaker accent for NZE participants for any story pairs, Tyneside participants were more accurate for TE speakers than NZE speakers, both in related story pairs, $\beta = 4.4$, $SE = 1.9$, $p < .001$, and unrelated story pairs, $\beta = 2.4$, $SE = 1.0$, $p = .038$. Moreover, the largest difference in key press accuracy was found for related story pairs, where Tynesiders' accuracy was 21.4 percentage points lower for the unfamiliar accent, as compared to a difference of 11.9 percentage points for unrelated story pairs.

Taken together, all participants did best for unchanged story pairs. Where story pairs included a change, Tyneside participants performed better with their own accent than for the NZE accent, especially when the changed words were semantically related. For New Zealand participants, there was a similar, albeit not significant, trend in the data. These findings suggest a familiarity benefit for change detection under adverse listening conditions. The increased task difficulty induced by presenting stories in noise might have been one reason why an effect of L1 accent was found here, but not in previous research [12].

3.2. Correction Accuracy

Out of the 955 change trials with a correct key press response, 797 trials (83.5%) were correct and 158 (16.5%) were incorrect as regards target recall (both original and changed word recalled correctly). Story pair was the only robust predictor of correction accuracy, $\chi^2(2) = 6.54$, $p = .011$. The difference between related and unrelated story pairs is shown in Figure 2. Participants were better at correcting semantically related rather than unrelated changes, regardless of the accent they heard and where they were from. There was an accuracy difference of 6.6 percentage points between the story pairs.

Thus, semantic proximity is helpful for the correction of the changes, with no difference in target recall due to familiar versus unfamiliar L1 accents, here presented in noise. The only evidence for a familiarity-based difference comes from the change detection results, where there was a familiarity benefit for the Tyneside participants. Further research is needed to clarify the mechanism behind this benefit. While a less-detailed processing account could explain the change detection results, it is not useful for the data regarding recall of changed words: for the latter, semantically related changes were corrected more accurately for both accents. Likewise, if the pattern of results is underpinned by a lexical processing imbalance between familiar and unfamiliar L1 accents, the differential effects of semantic proximity remain to be clarified.

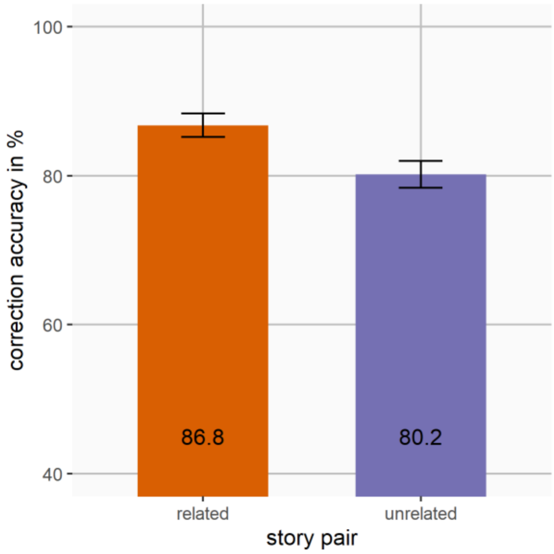

Figure 2: Target word correction accuracy by story pair (error bars: ± one SE)

Finally, auditory inspection by trained phoneticians determined that intelligibility was more greatly affected by noise masking for the NZE speakers than the TE speakers. This auditory impression was supported by long-term average spectra of the four speakers' voices, which showed that the energy of frequency components above 4 kHz was consistently lower for both NZE speakers. This might have artificially lowered the recall of NZE speech and, thus, the interaction between accent and participant origin could have been stronger if different speakers had been used. Importantly, this would not change the primary interpretation of the results because all participants were exposed to both NZE speakers, and the effects described here would likely remain significant.

4. CONCLUSION

This study used a change detection paradigm to investigate the effects of L1 accent familiarity on narrative recall in noise. The results showed a familiarity benefit for change detection for Tyneside participants, but not for New Zealand participants. This familiarity benefit was greater for Tynesiders' detection of semantically related story changes than for unrelated changes. While the exact mechanism behind these findings remains to be determined, it is clear that accent familiarity effects on narrative processing are not only mediated by the native language of the speaker: L1 accents, as well as L2 accents, can affect memory for spoken narratives.

5. ACKNOWLEDGEMENTS

The study presented here is part of the first author's doctoral research, which is funded by the Economics and Social Research Council (ESRC) in the United Kingdom (grant number: 2235617). We would like to thank Donald Derrick and Jen Hay at the University of Canterbury for their help with sourcing and recording the NZE speakers. We would also like to thank the speakers that lend their voices to this research and all participants that took part in it.

6. REFERENCES

[1] Lev-Ari, S., Keysar. B. 2012. Less-detailed representation of non-native language: Why non-native speakers' stories seem more vague. *Discourse Processes*, 49, 7, 523–538.
[2] Ferreira, F., Bailey, K. G. B., Ferraro, V. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 1, 11–15.
[3] Ferreira, F., Patson, N. D. 2007. The 'good-enough' approach to language comprehension. *Language and Linguistics Compass*, 1, 1–2, 71–83.
[4] Sturt, P., Sanford, A. J., Stewart, A., Dawydiak, E. 2004. Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review*, 11, 5, 882–888.
[5] Goslin, J., Duffy, H., Floccia, C. 2012. An ERP investigation of regional and foreign accent processing. *Brain and Language*, 122, 2, 92–102.
[6] Hanulíková, A., van Alphen, P. M., van Goch, M., Weber, A. 2012. When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, 24, 4, 878–887.
[7] Montgomery, C. 2018. The perceptual dialectology of England. In: Braber, N., Jansen, S. (eds.), *Sociolinguistics in England.* Palgrave, 127–164.
[8] Sharma, D., Levon, E., Yang, Y. 2022. 50 years of British accent bias: Stability and lifespan change in attitudes to accents. *English World-Wide*, 43, 2, 135–166.
[9] Adank, P., Evans, B. G., Stuart-Smith, J., Scott, S. K. 2009. Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 2, 520–529.
[10] Impe, L., Geerarts, D., Speelman, D. 2008. Mutual intelligibility of standard and regional Dutch language varieties. *International Journal of Humanities and Arts Computing*, 2, 1–2, 101–117.
[11] Sumner, M., Samuel, A. G. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60, 4, 487–501.
[12] Frances, C., Costa, A., Baus, C. 2018. On the effects of regional accents on memory and credibility. *Acta Psychologica*, 186, 63–70.
[13] Newcastle University IT Service. 2022. Newcastle University Form Builder. https://services.ncl.ac.uk/itservice/collaboration-services/form-builder/.
[14] Burridge, K., Kortmann, B. 2004. Introduction: Varieties of English in the Pacific and Australasia. In: Schneider, E., Burridge, K., Kortmann, B., Mesthrie, R., Upton, C. (eds.), *A handbook of varieties of English. Volume 1: Phonology*. De Gruyter, 567–579.
[15] Simion, D. 2022. Sound Bible: Party crowd. https://soundbible.com/2163-Party-Crowd.html.
[16] McCloy, D. 2013. Praat script: Noise masking of stimuli. https://github.com/drammock/praat-semiauto/blob/master/MixSpeechNoise.praat.
[17] Finger, H., Goeke, C., Diekamp, D., Standvoß, K., König, P. 2017. LabVanced: A unified JavaScript framework for online studies. In: *2017 International Conference on Computational Social Science IC^{2}S^2, Cologne, Germany, 10–13 July, 2016.*
[18] R Core Team. 2022. R: A language and environment for statistical computing. https://www.R-project.org/.
[19] Wickham, H. et al. 2019. Welcome to the Tidyverse. *Journal of Open Source Software*, 4, 43, 1–6.
[20] Bates, D. et al. 2022. lme4: Linear mixed-effects models using 'Eigen' and S4. https://cran.r-project.org/web/packages/lme4/index.html.
[21] Singman, H. et al. 2022. afex: Analysis of factorial experiments. https://cran.r-project.org/web/packages/afex/index.html.
[22] Lenth, R. V. et al. 2022. emmeans: Estimated marginal means, aka least-squares means. https://cran.r-project.org/web/packages/emmeans/index.html.