The Impact of Cognitive Load on Speech Production in German-English Bilinguals

Jessica Göbel[1,2], Kathleen M. McCarthy[2]

[1] New York University and [2]Queen Mary University of London
jessica.goebel@nyu.edu, k.mccarthy@qmul.ac.uk

ABSTRACT

This study investigates cognitive load effects on L1 transfer and style-shifting in ten German L1-English L2 bilinguals. Participants were asked to complete two speech tasks, reading and recalling a text, under varying cognitive load conditions. The lower load condition consisted of speakers simply reading or recalling a text out loud. The higher load condition included a distractor task whereby speakers were instructed to read or recall a text out loud while simultaneously concentrating on arithmetical calculations. Both variables examined – word-final light /l/ and word-final devoiced /d/ – showed a very consistent trend towards an increase in the German-accented variant in the greater load condition. In other words, speakers appear to default to a more typical German-accented style when attention is diverted. Few studies have explicitly examined cognitive load effects on bilingual speech production, thus, this paper sheds light onto the cognitive demands that influence linguistic variability for multilinguals.

Keywords: cognitive load, language transfer, bilingualism.

1. BACKGROUND

Language transfer is a relatively well-studied phenomenon in bilingual speech research [9]. However, few studies have explored how fluctuations in cognitive demands can impact the speech production of bilinguals. We know that divided attention, for example, can interrupt the speech processes of monolinguals. Sharma and McCarthy [18] found that an increase in cognitive load can lead to an increase in vernacular features. Furthermore, previous sociophonetic research on monolinguals has found that an increased cognitive load can disrupt speech convergence processes [1; see e.g., 16 for a review]. The current study asks how attentional multi-tasking affects bilingual speech, specifically how it affects the type and rate of L1 transfer in formal speech. Our research question is: Does an increase in cognitive load result in an increase in L1 transfer?

Language transfer is when a linguistic feature is applied from one language to another by a multilingual speaker [9]. Forward transfer, the process whereby a speakers' first language (L1) influences their second language (henceforth L1 transfer) is well-attested in speech production research [e.g., 9; 10]. Multiple factors, such as age of acquisition, proficiency, or the structural differences between the L1 and L2, have been shown to influence the extent of L1 transfer [2; 8].

Another potential factor at play is formality. Passoni et al. [14] found that formality predicted pitch range in Japanese L1-English L2 bilinguals. Bilinguals significantly varied their pitch range according to level of formality in Japanese, but not English. The authors explained this finding in terms of indexical differences between the two languages: pitch is used to index politeness and formality in Japanese but not in English. The current paper aims to build on Passoni et al. [14] by examining bilinguals across formal read speech and casual interview speech, thus furthering understanding of the formality association of non-native accented English.

The role of formality on speech-style has been thoroughly studied in style-shifting research. According to Labov's [11] attention-paid-to-speech model, casual conversational speech has the lowest attention paid to speech while formal read speech has the highest attention paid to speech. Specifically, in casual conversations, speakers tend to revert to their first learned speech style [12]. Importantly, Labov's style shifting model does not unambiguously differentiate between social and cognitive factors that could affect style shifting as interview speech differs from read speech both socially and cognitively.

Recent work on style control has found that cognitive factors, such as cognitive load, might play a key role in style shifting. Sharma and McCarthy [18] examined the effect of cognitive load on style control in monolinguals. They employed a design that allowed them to test the impact of cognitive load while eliminating contextual differences across speech tasks by instructing participants to use a formal news report style throughout. The study found that an increase in cognitive load led to higher rates of vernacular features, thus inhibiting the speakers' ability to maintain their potentially later learned style.

For bilingual speakers, increasing the cognitive load when speaking their L2 could result in an increase in L1 transfer. Sorace [19] found that sentence ambiguities affect L2 syntactic processing. Specifically, sentence ambiguities led to a greater than normal processing load and induced L2 learners into shallow processing, resulting in more syntactic errors. At the phonetic level, Pajak et al. [13] found that that L2 learners have difficulty producing similar-sounding words. They propose that producing highly similar sounding words with novel sound contrasts not present in speakers' L1 increases the cognitive load. This in turn leads to higher rates of non-target-like productions.

The above literature suggests that increased cognitive load impacts language production and perception for bilinguals as it does for monolinguals. However, studies which examine cognitive load in a controlled experimental setting for bilinguals are lacking. This underscores the need for the present paper. In the current study we focused on the speech of German L1-English L2 bilinguals. Based on previous findings, we hypothesize that a higher cognitive load would lead to significantly greater rates of German-accented realisations and L1 transfer. Ten German L1-English L2 bilinguals were asked to complete two speech tasks, reading and recalling a text, under varying cognitive load conditions. The primary aim is to examine cognitive load effects on rates of German-accentedness in order to shed light on how cognitive load can influence language transfer. A secondary aim is investigating how level of formality affects rates of German-accentedness.

Two variables believed to be sensitive to transfer were considered: word-final light /l/ and word-final devoiced /d/. Standard Southern British English has a sharp light /l/ dark /l/ contrast, whereby /l/ is realised as [ɫ] word-finally and as [l] word-initially [21]. In comparison, Standard German lacks this light /l/ dark /l/ contrast. In Standard German, /l/ is always realised as light, making this variable ideal for an analysis of L1 transfer [17].

The second variable, word-final /d/, varies across English and German. Final obstruent devoicing is a highly regular phonological process that takes place in German and does not exist in English. Özaslan and Gabriel [23] find that final obstruent devoicing is transferred from German to English by L2 learners, yielding high rates of non-target-like productions. Although final obstruent devoicing applies to plosives, fricatives, and affricates, the current study will limit its focus to word-final /d/.

2. METHOD

2.1 Participants

The speakers are ten German L1-English L2 bilinguals between the ages of 20 and 29. All are Standard German speakers who considered German to be their first and dominant language. The participants learned English as an L2 in elementary school (except for one speaker who learned English before elementary school, but not from birth). At the time of recording, the speakers lived in Germany or moved out of Germany relatively recently. Age of English acquisition varied from 3 to 12, with an average age of English acquisition of 8.5 years**.**

2.2. Materials and procedure

The experimental design follows Sharma and McCarthy's [18] study on cognitive load and style shifting in monolingual speakers. To maintain the same speech style across the different tasks and cognitive load conditions, participants were instructed to speak as if they were presenting a radio news report across the different tasks. Speakers were given a practice report to read aloud at the start of the experiment to enable them to establish their newsreader style.

The experiment consisted of two tasks. Task 1 (Read Speech) involved reading a text out loud in a formal newsreader style. The task consisted of two conditions. Condition A (lower cognitive load) consisted of speakers simply reading a text out loud. Condition B (higher cognitive load) included a distractor task where speakers were asked to read a text out loud and simultaneously concentrate on an audio stream of numbers (inter-stimulus interval [ISI]: 2 seconds). Speakers were instructed to add up the numbers while they read and report the total sum after giving their report [see 22 for this design].

Task 2 (Recall Speech) consisted of reading a news report and being asked to memorize it as best as they could before giving the report from memory. This task also involved two cognitive load conditions. Condition A (lower cognitive load) simply involved speakers reading, memorizing, and then recalling the text out loud. Condition B (higher cognitive load) also used a distractor task where speakers were asked to read, memorize, and then recall a text while simultaneously concentrating on an audio stream of numbers (ISI: 2 seconds). Speakers were asked to report the first five numbers they heard after completing their report [see 7 for this design]. Following Sharma and McCarthy [18], an easier distractor task was used for recall speech as simultaneous arithmetical calculations were too difficult due to working memory effects of memorizing a text. Out of ten participants, two speakers correctly solved one of the number tasks in the higher cognitive load condition.

To control for confounding effects of cognitive load order of the tasks, the order of the cognitive load conditions (high or low) was counterbalanced across speakers. To control for confounding effects of task type (i.e, read vs recall), the order of the read and recall speech was counter-balanced across speakers.

After the experimental component was complete, the first author conducted casual sociolinguistic interviews with participants to gather information about speakers' subjective responses to the different tasks, their language background, as well as getting a baseline of casual speech. The interviews and experiment were conducted remotely using Zoom and speakers were asked to record themselves on their phone to ensure that internet connectivity issues would not affect the audio quality. Gorilla Experiment Builder [3] was used to design the experiment so that participants could read and recall texts and listen to the audio distractor at the same time. The interviews were conducted in English, though all speakers knew that the interviewer was a German-English bilingual.

2.3 Scripts and variables

All scripts were adapted from Sharma and McCarthy [18]. The reading scripts were on science news stories to target a formal style but avoid political material that may prompt affect. Texts were edited to include tokens for three variables thought to be sensitive to L1 transfer: word-final /l/ and /d/. Care was taken to control for clausal context, lexical frequency, and phonetic context in the texts as much as possible. Following and preceding phonological contexts of both variables were additionally considered in statistical analyses. Variables were coded auditorily by the first author, consulting the spectrogram in Praat [6] for ambiguous tokens. A random sample of 20% of the tokens were coded by the second author and yielded an 84% inter-rater agreement.

For ambiguous /l/ tokens, the F1 and F2 contours were considered. If the F1 was relatively high accompanied by a very low F2 towards the end of the lateral, the token was classed as dark. If the F1 and F2 contour remained relatively flat, the token was classed as light. For the analysis of word-final /l/ – whether it is realized as dark [ɫ] or light [l] – only non-cluster contexts were considered, following [8]. This meant that word final /l/ could be followed by a non-/l/ consonant, vowel in the following word, or pause.

For /d/, the presence (or absence) of a continuous voice bar determined whether the item was coded as voiced (or devoiced). Coding for word-final /d/ – whether voiced [d] or devoiced [t] – also followed previous work [23]. Word-final and word medial contexts were considered, and the following phonological environments included non-/d, t/ consonants, vowels, or pauses. Preceding contexts included voiced consonants and vowels.

3. RESULTS

To analyse which constraints are the most accurate predictors for L1 transfer, mixed-effect logistic regression models were created in R [15] using the lme4 package [5]. The models included cognitive load (high or low) and task type (read, recall, or casual speech) as fixed factors, and the speaker and word as random intercepts.

3.1 Overall distribution

Table 1 shows the overall proportion of variants for all 10 speakers, across all conditions. In total, 970 tokens were collected, an average of around 50 tokens per speaker per variable. Both variables show nearly a 50-50 split between variants, with rates of word-final light /l/ at 49%, and rates of word-final devoiced /d/ at 54%, which suggests a high degree of variability among speakers.

	Word-final /l/		**Word-final /d/**	
Variant	light*	dark	devoiced*	voiced
N	212	217	292	249
%	49%	51%	54%	46%

Table 1. Overall distribution of all variants: German-accented English* vs. British English realisations

3.2 Cognitive load and task effects

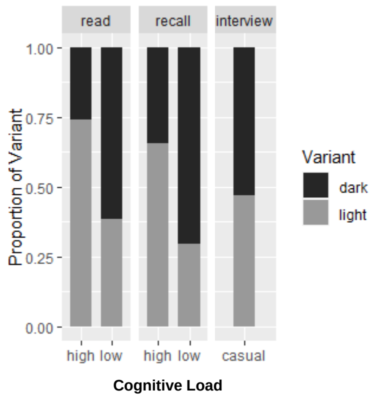

Figure 1. Rates of word-final light /l/ according to task

Figure 1 shows the proportion of word-final light /l/ (the German-accented variant) according to task type and cognitive load. Greater cognitive load elicited higher rates of German-accented /l/ in both read and recall speech than the lower cognitive load. Additionally, casual speech (interview) had higher rates of the German-accented variant than read and recall speech in the low cognitive load condition. A main effect of cognitive load (Std. Error = 0.428; z = -4.402; p < 0.001) and task type (Std. Error = 0.472; z-value = -2.310; p = 0.021) was found, whereby interview speech had significantly higher rates of the German-accented variant than read speech. The interaction between task type and cognitive load (Std. Error = 0.722; z-value = -0.441; p = 0.659) was not significant.

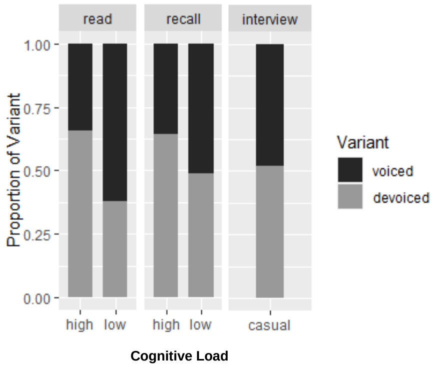

Figure 2. Rates of word-final devoiced /d/ according to task

Figure 2 shows the proportion of word-final devoiced /d/ (the German-accented variant) according to task type and cognitive load condition. Greater cognitive load led to higher rates of the German-accented variant. Additionally, we can see that casual interview speech had the highest rates of German-accented /d/, while formal read speech (in the low load condition) had the lowest rates of German-accented /d/, with recall speech patterning somewhere in between. The mixed model found a main effect of cognitive load (Std. Error = 0.439; z-value = -3.881; p < 0.001) and task type (Std. Error = 0.423; z-value = -2.800; p = 0.005), whereby interview speech had significantly higher rates of the German-accented variant than read speech. The interaction between task type and cognitive load was not significant (Std. Error = 0.556; z-value = 1.251; p = 0.211).

4. DISCUSSION

The main aim of our paper was to examine cognitive load effects on L1 transfer. Both variables show the same cognitive load pattern – greater load conditions led to significantly higher rates of the German-accented variant. As task type and cognitive load were fully controlled, the findings illustrate that speakers were more likely to default to a more typical German-accented style when attention is diverted. A reversion to vernacularity or a first-learned speech style in higher cognitive load conditions has been reported for monolinguals [e.g. 11; 18]. However, to our knowledge, cognitive load has never been examined in a bilingual context in a controlled experimental setting. Thus, this paper furthers understanding of how cognitive load affects L2 speech production and shows that diverted attention can lead to an increase in L1 transfer and cause speakers to default to a more German-accented style. This is potentially due to the cognitive primacy of speakers' first language [11]. Note, however, that the small participant sample renders these results less statistically robust.

Some speakers informally reported that they felt their speech was more German-accented in the greater cognitive load condition. One speaker noted that "in the last reading task it was less of a German accent and the first one was more just because […] you need to focus on the numbers and reading and also […] on your accent, so this is three things at the same time." This provides tentative evidence that some speakers have an awareness of their varying rates of German-accentedness and presents some qualitative evidence for the quantitative patterns observed.

A secondary aim of the current study was to examine the formality-associations of German-accented speech through a consideration of both casual interview and formal newscaster speech. We found that casual interview speech elicited significantly higher rates of German-accented /d/ and /l/ than formal read or recall speech. This finding suggests that German-accented speech is considered less formal than certain native varieties of English. The finding that word-final light /l/ and devoiced /d/ are sensitive to task type (read or interview) provides at least some evidence that rates of German-accentedness is somewhat driven by formality-associations. This finding is not altogether surprising as previous research has found that people associate non-native accents with a lack of linguistic competence and a lack of formality [4; 20].

The effect of task type (read, recall, or interview) and cognitive load (high or low) indicates that rates of L1 transfer are both socially and cognitively driven, whereby rates of German-accentedness increase in casual speech and in higher cognitive load conditions. Note that the effect of cognitive load was stronger than the effect of formality.

In sum, our findings suggest that cognitive factors may play a role in the variation of L2 speech production patterns and should therefore be considered alongside other well-researched social factors (e.g., motivation, attitudes) when conducting multilingual speech research.

5. REFERENCES

[1] Abel, J.,Babel, M. 2017. Cognitive load reduces perceived linguistic convergence between dyads. *Language and Speech* 20, 479–502.

[2] Abrahamsson, N., Hyltenstam, K. 2009. Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(1), 249–306.

[3] Anwyl-Irvine, A., Massonié, J., Flitton, A., Kirkham, N., Evershed, J. 2019. Gorilla in our midst: an online behavioural experiment builder. Behavior Research Methods. Doi: https://doi.org/10.3758/s13428-019-01237-x

[4] Ballard, L., Winke, P. 2017. Students' Attitudes Towards English Teachers' Accents: The Interplay of Accent Familiarity, Comprehensibility, Intelligibility, Perceived Native Speaker Status, and Acceptability as a Teacher. In T. Isaacs & P. Trofimovich (Eds.), *Second Language Pronunciation Assessment: Interdisciplinary Perspectives* 1(107), 121–140.

[5] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48. doi:10.18637/jss.v067.i01.

[6] Boersma, P, Weenink, D. 2021. Praat: doing phonetics by computer [Computer program]. Version 6.1.40, retrieved 27 February 2021 from http://www.praat.org/.

[7] Christodoulides, G. 2016. Effects of Cognitive Load on Speech Production and Perception. Doctoral dissertation, Université catholique de Louvain.

[8] de Leeuw, E., Menne, I., Scobbie, J. 2012. Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism* 17(6), 683–700.

[9] Flege, J. 1995. Second-language speech learning: Theory, findings, and problems. In: Strange, W. (ed.) *Speech perception and linguistic experience: Issues in cross-language research*, 229–273. Timonium, MD: York Press.

[10] Kim, K., Joo, K. 2021. L1 transfer in the interpretation of L2 reflexive pronouns by child learners of Korean. *International Journal of Bilingualism* 25(6), 1529–1544.

[11] Labov, W. 1966. *The Social Stratification of English in New York C*ity. Cambridge: Cambridge University Press.

[12] Labov, William. 2013. The Language of Life and Death: The Transformation of Experience in Oral Narrative. Cambridge: Cambridge University Press.

[13] Pajak, B., Creel, S. C., Levy, R. 2016. Difficulty in learning similar-sounding words: A developmental stage or a general property of learning?. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(9), 1377.

[14] Passoni, E., de Leeuw, E., Levon, E. (2022). Bilinguals Produce Pitch Range Differently in Their Two Languages to Convey Social Meaning. *Language and Speech*, *65*(4), 1071–1095. https://doi.org/10.1177/00238309221105210

[15] RStudio Team. 2019. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

[16] Roelofs, A., Piai, V. 2011. Attention demands of spoken word planning: A review. *Frontiers in psychology* (2), 1–14.

[17] Recasens, D. 1995. Velarization degree and coarticulatory resistance for /l/ in Catalan and German. *Journal of Phonetics*, 23, 37–52.

[18] Sharma, D., McCarthy, K. 2018. Attentional load and style control. *University of Pennsylvania Working Papers in Linguistics* 24(2), 127–136.

[19] Sorace, A. 2006. Possible manifestations of shallow processing in advanced second language speakers. *Applied Psycholinguistics*, 27, 88–91.

[20] Subtirelu, N. 2015. "She does have an accent but...": Race and language ideology in students' evaluations of mathematics instructors on RateMyProfessors.com. *Language in Society* 44, 35–62.

[21] Wells, John Christopher. 1982. *Accents of English* (Vols.1&2). Cambridge: Cambridge University press.

[22] Yin, B., Ruiz, N., Chen, F., Khawaja, A. 2007. Automatic cognitive load detection from speech features. In Proceedings of OzCHI 2007.

[23] Özaslan, M., Gabriel, C. 2019. Final obstruent devoicing in English and French as foreign languages: Comparing monolingual German and bilingual Turkish-German learners. In S. Thiele & C. Gabriel (Ed), *Romanische Sprachen in ihrer Vielfalt: Brückenschläge zwischen linguistischer Theoriebildung und Fremdsprachenunterricht*, 177–209. Stuttgart: Ibidem.