Attention, Perception, & Psychophysics (2020) 82:1558–1565
https://doi.org/10.3758/s13414-019-01946-7

SHORT REPORT

When processing costs impact predictive processing: The case of foreign-accented speech and accent experience

Vincent Porretta[1] · Lori Buchanan[1] · Juhani Järvikivi[2]

Published online: 22 January 2020

Abstract
Listeners use linguistic information and real-world knowledge to predict upcoming spoken words. However, studies of predictive processing have focused on prediction under optimal listening conditions. We examined the effect of foreign-accented speech on predictive processing. Furthermore, we investigated whether accent-specific experience facilitates predictive processing. Using the visual world paradigm, we demonstrated that although the presence of an accent impedes predictive processing, it does not preclude it. We further showed that as listener experience increases, predictive processing for accented speech increases and begins to approximate the pattern seen for native speech. These results speak to the limitation of the processing resources that must be allocated, leading to a trade-off when listeners are faced with increased uncertainty and more effortful recognition due to a foreign accent.

Keywords Foreign-accented speech · Prediction · Language experience · Visual world paradigm · Eyetracking

Substantial research has indicated that language processing is inherently predictive, with language users processing the current input while simultaneously anticipating upcoming information on the basis of prior experience/input (see Clark, 2013, for an overview). Prediction of the most likely content and structures enhances the accuracy and efficiency of comprehension (e.g., Altmann & Kamide, 1999; Trueswell, Tanenhaus, & Kello, 1993). Prediction is believed to be a graded phenomenon that is probabilistic rather than all-or-nothing (Kuperberg & Jaeger, 2016), with cues working in tandem (Henry, Hopp, & Jackson, 2017). Predictive processing influences both visual (e.g., Rayner, Slattery, Drieghe, & Liversedge, 2011) and spoken (e.g., Kamide, 2008) language processing. Evidence has come primarily from eyetracking (e.g., Rayner et al., 2011, for reading; and Altmann & Kamide, 1999, for the visual world paradigm, VWP), and event-related potentials (ERP; e.g., Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Both semantic (e.g., Altmann & Kamide, 1999) and morphosyntactic (e.g., Huettig & Janse, 2016) types of information trigger predictions, and these processes are believed to be fundamental and automatic (Clark, 2013). However, individual differences can mediate prediction (Huettig, 2015). Studies of prediction during spoken language processing have primarily focused on predictive processing under optimal listening conditions. However, in everyday conversations the acoustic signal that guides prediction is remarkably varied, due to external noise and speaker variability. Here we focus on interspeaker variability—specifically, the uncertainty introduced by a foreign accent.

The presence of a foreign accent typically results in processing costs (e.g., Bradlow & Bent, 2008; Porretta, Tucker, & Järvikivi, 2016), which are thought to arise due to the dynamics of lexical activation. Porretta et al. (2016) found that as accent strength increases, spoken primes become less effective, indicating reduced activation. Additionally, Porretta and Kyröläinen (2019) demonstrated that foreign-accented speech induces more lexical competition; listeners entertain more candidate words for a longer period of time, even when comprehension is successful. The uncertainty of the signal likely leads to these changes, as similar results have been found for speech in noise (Brouwer & Bradlow, 2016). At the same time, these effects are ameliorated by long-term experience

✉ Vincent Porretta
porretta@uwindsor.ca

[1] University of Windsor, Windsor, ON, Canada
[2] University of Alberta, Edmonton, AB, Canada

with accented speech (e.g., Bradlow & Bent, 2008; Porretta et al., 2016).

The extent to which foreign-accented speech impacts predictive processing and how accent experience may modulate this effect remain unclear. Two studies (Goslin, Duffy, & Floccia, 2012; Romero-Rivas, Martin, & Costa, 2016) have provided partial, and also conflicting, evidence of the impact of accented speech, by examining the N400 ERP component. Romero-Rivas et al. demonstrated that there was no difference between accented and unaccented conditions with regard to lexical preactivation of best-fitting and unrelated words in a sentence context. Goslin et al., in contrast, demonstrated that accented low-cloze-probability words elicited a reduced N400 as compared to an unaccented condition. However, neither study examined the potential effect of accent experience.

Anticipatory eye movements are contingent on predictions about forthcoming information and can be harnessed in order to examine predictive processing. To investigate whether foreign-accented speech influences predictive processing, and whether lifelong accent experience mediates this process, we replicated and extended the seminal study by Altmann and Kamide (1999), which utilized anticipatory eye movements as an indicator of predictive processing. Using VWP eyetracking, they presented participants with visual scenes containing multiple objects (e.g., ball, cake, car, train) and spoken sentences (e.g., *the boy will move/eat the cake*). The verb either selected all four objects (*move* condition), or only one (*eat* condition). In the restricting (i.e., *eat*) condition, participants were more likely to look at the cake prior to hearing the word *cake*. This study showed that information at the verb restricts the reference of a yet-unencountered grammatical object, indicating a predictive relationship between verbs, syntactic objects, and visual context. Here, similar methods and materials examine anticipatory eye movements when processing foreign-accented speech. We expected that, although semantic constraints facilitate prediction, the presence of an accent would reduce this benefit. We further expected that lifelong experience with the accent would modulate the impediment imposed by accented speech, such that greater accent experience would lead to enhanced prediction.

Method

Participants

Sixty native speakers of English (47 female, 13 male) were recruited from the University of Windsor (18–40 years of age, M = 21.17, SD = 3.84).[1] All participants reported normal (or corrected-to-normal) vision and normal hearing. In accordance with approval from the University of Windsor Research Ethics Board, the participants provided written informed consent and received partial course credit.

Stimuli

Following Altmann and Kamide (1999), the critical stimuli consisted of simple English transitive sentence pairs (N = 24), such that the verb either restricted or did not restrict the direct object—for example, *The fireman will climb the ladder*, in which *climb* restricts the type of object (*ladder*) that can follow it, versus *The fireman will need the ladder*, in which *need* could select for many different objects. Additionally, 24 simple transitive clause filler sentences were created. All of the stimuli were produced by one male native speaker of English and one male native speaker of Mandarin Chinese. All sentences were normalized for amplitude. The mean duration was 1,934 ms (SD = 226) for the native sentences, and 2,380 ms (SD = 296) for the nonnative sentences. For the native talker, the mean duration between verb onset and object onset was 530 ms (SD = 89.4), and the mean duration of the target object was 576 ms (SD = 112). For the nonnative talker, the mean duration between verb onset and object onset was 638 ms (SD = 107), and the mean duration of the target object was 696 ms (SD = 152).[2]

The critical and filler items were presented to 36 (32 female, four male) University of Windsor students[3] 18–57 years of age (M = 22.44, SD = 7.16) in a separate transcription/rating task. The raters listened to each sentence over headphones, with no other information about the talkers or sentences, completing one of four counterbalanced lists. Each sentence was first transcribed and then rated on a scale from 1 (*no foreign accent*) to 9 (*very strong foreign accent*). The transcriptions were scored for keyword intelligibility—that is, the combined accuracy of the subject, verb, and object. The talkers differed significantly (see Table 1) in both mean intelligibility and mean accentedness. However, although the nonnative talker had a moderately strong accent, he was highly intelligible.

The spoken stimuli were paired with visual arrays containing the subject (e.g., *fireman*) in the center, surrounded by four equidistantly placed object images. For the critical stimuli, these included the target object (e.g., *ladder*) and three other objects (e.g., *hose*, *axe*, *paperclip*) that grammatically completed the sentence and were semantically plausible completions for the nonrestricting context (e.g., when the verb is *need*). For the filler sentences, the four

[1] Twenty-seven of the participants reported being bilingual, though none in any variety of Chinese.

[2] All duration comparisons between the native and nonnative talkers were significant ($ps < .0001$).

[3] These participants did not take part in the eyetracking study.

Table 1. Summary of talker intelligibility and accentedness

Measure	Mean (SD)			
	Native talker	Non-native talker	*t* value	*p* value
Intelligibility	.99 (.02)	.94 (.03)	10.26 (35)	< .0001
Accentedness	1.07 (0.1)	6.47 (1.58)	– 20.84 (35)	< .0001

objects had the same properties; however, the target object was not depicted on screen. For half of the fillers, the objects mimicked the nonrestricting context (i.e., were semantically plausible completions); for the other half, the objects mimicked the restricting context (i.e., were not semantically plausible completions).

The black-and-white images were selected from various sources of standardized pictures, including the Snodgrass and Vanderwart (1980) picture set, the International Picture Naming Project (Szekely et al., 2004), and the Bank of Standardized Stimuli (Brodeur, Guérard, & Bouras, 2014). For objects not found in the databases, similarly styled, freely available online drawings were selected. Each image was only seen once during the experiment. The stimuli included in this study (i.e., audio files, sentences, object sets, and image IDs) are available via the Open Science Framework.[4]

Procedures

Participants sat at a chinrest situated in front of a desktop-mounted EyeLink 1000 Plus eyetracker (SR Research Ltd.) recording at 1000 Hz. The system was calibrated to the participant's right eye using a 9-point calibration procedure. Sentences were presented over speakers, and image arrays were displayed on screen. The subject was always presented in the center of each array, with the object images at equidistant locations. The target position was balanced across trials. Written instructions were provided along with two practice items. Participants were presented with one of four counterbalanced lists, such that each sentence was presented in one of the four conditions (i.e., native nonrestricting, native restricting, nonnative nonrestricting, nonnative restricting). Items were blocked by talker, the block order was randomized, and items were randomized within blocks.

Each trial began with a 500-ms central fixation cross, followed by the visual array. After 200 ms the auditory stimulus was presented (see McQueen & Viebahn, 2007), and participants indicated via button press whether the visual array matched the auditory sentence. Subsequently, participants responded to a brief questionnaire (see Table 4 in the Appendix) designed to estimate their lifetime experience interacting with Chinese-accented speakers.

[4] https://osf.io/2r3zc/?view_only=8697f901a91e453e957dfdb12e230de1

Data preparation

Time-series sample data (relative to object word onset) were processed using the R package VWPre (Porretta, Kyröläinen, van Rij, & Järvikivi, 2016). Using 50-ms windows within each recording event, the proportions of samples falling within and outside each interest area were calculated and converted to empirical logits with variance weights (see Barr, 2008). The picture verification responses indicated that participants were highly accurate ($M = .94$, $SD = .05$, range = .77–1), with only 186 errors in 2,880 trials. Incorrect trials (6.46% of the data) were removed prior to the analysis. As is shown in Fig. 1, looks to the target object occurred earlier in the restricting than in the nonrestricting conditions for both talkers.

Statistical considerations

Generalized additive mixed modeling (mgcv, version 1.8-24; Wood, 2018) was used to analyze the time-series data in R. Predictors and interactions were evaluated by the estimated *p* value of the smoothing parameter/parametric component and maximum likelihood (ML) score comparison of model variants. Delta AIC (ΔAIC; Akaike information criterion [AIC] of a simpler model minus AIC of a more complex model) was used to evaluate the strength of evidence for critical interactions, by means of information loss (Burnham & Anderson, 2002). In general, a ΔAIC less than 2 suggests substantial evidence for the simpler model; a ΔAIC between 3 and 7 indicates considerably less support for the simpler model; and a ΔAIC greater than 10 indicates that the simpler model is very unlikely (Burnham & Anderson, 2002). Effects and differences were calculated using itsadug (van Rij, Wieling, Baayen, & van Rijn, 2017).

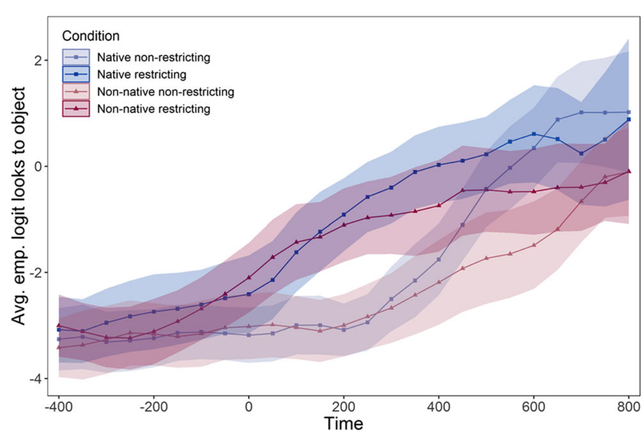

Fig. 1. Average looks to the target by condition, with 95% simultaneous confidence bands. The critical time points are – 400 (average acoustic onset of the verb), 0 (acoustic onset of the target object), and 800 (average acoustic offset of the target object).

Results

Analysis 1: Accent and prediction

The dependent variable was the empirical logit of looks to the target object image from – 400 ms to 800 ms (relative to the acoustic onset of the target object word), representing the likelihood of looking at the target between the average acoustic onset of the verb and the average acoustic offset of the object. The primary independent variable of interest was condition, the combination of talker (native vs. nonnative) and verb type (restricting vs. nonrestricting), which was treatment-coded with *native-nonrestricting* as the reference level. Because the time course of processing was of critical interest, time (in milliseconds) was included as a covariate. Trial order and log frequency of the target object (from the English Lexicon Project; Balota et al., 2007) were included as control variables.

The model was fitted with by-subject and by-item factor smooths for time and by-event random intercepts. Factor smooths allow the shape of the average time course to vary by participant and item. Random intercepts for event (the combination of subject and trial) allow a unique intercept for each time series. For frequency and trial, nonlinear functional relations for the response variable—smooth functions (Baayen, 2010; Wood, 2017)—were entered. Condition and time were included as a smooth interaction, with condition set as a parametric component. Because autocorrelation in time-series data can lead to overconfidence (Baayen, van Rij, Cecile, & Wood, 2018), an AR-1 correlation parameter, ρ = .74, estimated from the data, was included.[5] Finally, we included the inverse of the empirical logit variance estimates as weights in the model (see Barr, 2008). The model was then trimmed (see Baayen, 2008), removing 13 data points (0.04%). ΔAIC (168.5) indicated substantial support for the interaction between time and condition. The results of the model are presented in Table 2.

The conditional smooths for time resulted in significantly difference curves. These differences are displayed in Fig. 2 and correspond to the difference between similarly colored lines in Fig. 1. The difference curves are presented by talker, indicating more looks to the target in the restricting condition (i.e., *climb the ladder*) than in the nonrestricting condition (i.e., *need the ladder*) for both talkers. These two curves are different from one another when neither lies within the other's confidence band (see the nonshaded portions of Fig. 2). The first significant period persisted for 121 ms, from the average onset of the verb (– 400 ms) until 279 ms before the onset of the target object. The second persisted for 448 ms, beginning 36 ms prior to the onset of the target object and ending 412 ms after that onset.

[5] Because factor smooths for time can improve autocorrelated residuals, ρ was determined after fitting the random-effects structure.

Analysis 2: Prediction and experience

A second analysis was carried out only on trials presented in the nonnative voice, to examine the effect of accent experience.[6] The dependent variable was the same as in Analysis 1. The primary independent variable of interest was listener experience with Chinese-accented English, established via the questionnaire. Participants estimated their total lifetime experience interacting with speakers with a Chinese accent as a percentage of their lifetime interactions. The measure (range = 0–30, M = 7.04, SD = 6.54) contained a right skew. Following Porretta and Tucker (2019) and Porretta et al. (2016), log transformation (with a constant of 1) was employed (range = 0–3.43, M = 1.78, SD = 0.82).

The model was fitted as in Analysis 1, with the same random-effects structure and control variables. Time, experience, and verb type (restricting vs. nonrestricting) were included as a three-way interaction using a tensor product (see Wood, 2017). Verb type was set as a parametric component, and weights were included. The AR-1 correlation parameter was estimated to be ρ = .75. The model was then trimmed, removing 31 data points (0.18%). ΔAIC (103.9) indicated substantial support for verb type in interaction with both time and experience. Likewise, the ΔAIC (59.9) indicated substantial support for experience in interaction with time and verb type. The results of the model are presented in Table 3.

Figure 3 displays the significant interaction between time, accent experience, and verb type. In panel A (nonrestricting condition), we see no influence of experience prior to 200 ms, as expected; additionally, as expected from previous research (Porretta et al., 2016), experience influenced looks to the target as the critical object was being heard. By contrast, panel B (restricting condition) indicates that prior to 200 ms, participants with greater experience began looking at the target object earlier. Panel C represents the difference between the two conditions; the shaded mask indicates the regions that were not different. Participants with the lowest experience displayed no prediction in the first half of the time window, but participants in the mid-range of experience displayed a prediction effect. Strikingly, the participants with the greatest amount of accent experience showed the strongest prediction effect around 200 ms, which then diminished, mirroring the pattern seen in Fig. 2 for the native talker.

[6] An analogous analysis was carried out on the stimuli spoken in the native voice. The difference surface did not indicate an influence of experience on the prediction effect.

Table 2. Summary of the generalized additive mixed model for accent and prediction

	Estimate	*SE*	*t* Value	*p* Value
Parametric coefficients				
Intercept	– 0.86	0.16	– 5.47	< .0001
NonNative.NonRestricting	– 0.14	0.11	– 1.27	.2044
Native.Restricting	0.33	0.11	3.03	.0024
NonNative.Restricting	0.13	0.11	1.14	.2555
	E*df*	Ref. *df*	*F* Value	*p* Value
Smooth terms				
Smooth: Time, Native.NonRestricting	4.83	5.88	13.78	< .0001
Smooth: Time, NonNative.NonRestricting	7.36	8.27	6.40	< .0001
Smooth: Time, Native.Restricting	3.20	3.94	8.16	< .0001
Smooth: Time, NonNative.Restricting	1.02	1.03	11.51	.0006
Smooth: Trial	1.00	1.00	0.03	.8691
Smooth: Frequency	1.00	1.00	0.17	.6846
Random effect: Time, Subject	376.22	539.00	40.62	< .0001
Random effect: Time, Item	150.12	214.00	47.40	< .0001
Random effect: Event	1,079.08	1,388.00	4.66	< .0001

General discussion

The data indicate that a foreign accent interferes with predictive processing. In Analysis 1, we replicated the prediction effect reported by Altmann and Kamide (1999) for native-accented speech. Although there was some indication of prediction for foreign-accented speech, the magnitude of this effect was reduced (and delayed) relative to the native accent. Importantly, even when listening to foreign-accented speech, listeners appear always to predict to the extent possible. Thus, it does not appear that listeners simply "shut off" prediction in the presence of an accent, aligning with Kuperberg and Jaeger's (2016) assertion that predictive processing is graded. The present results can be explained by two related but alternative views of prediction.

Under the first view, the processing demands related to decoding accent-related variability prevent the full engagement of anticipatory processes. As a result of uncertainty in mapping the acoustic input to phonological categories, decoding requires more effort and time, influencing how the limited resources are allocated dynamically. Thus, fewer anticipatory eye movements would reflect lesser (or a total lack of) engagement of prediction. Under the second view,[7] prediction is always fully engaged, though the uncertainty of the decoding process is inherited by the prediction mechanism. Thus, fewer anticipatory eye movements would reflect making predictions from uncertain data.

Using pupil dilation, Porretta and Tucker (2019), showed that accented speech requires more listening effort. Additionally, Porretta and Kyröläinen (2019) demonstrated that accented speech results in the activation of more lexical competitors, which creates more possibilities and requires more time to resolve (if it is resolved at all). This could explain how the output of the decoding process influences the engagement of the prediction mechanism. Signal decoding requires additional effort and would take precedence over prediction. This is reasonable if prediction requires at least some degree of certainty of the input. In some cases, the process might take too long for prediction to be beneficial. However, it has also been shown that listeners maintain uncertainties in speech perception (Brown-Schmidt & Toscano, 2017), and specifically for accented speech (Burchill, Liu, & Jaeger, 2018). If uncertainty is maintained, then uncertain input would lead to uncertain predictions, which might not warrant eye movement.

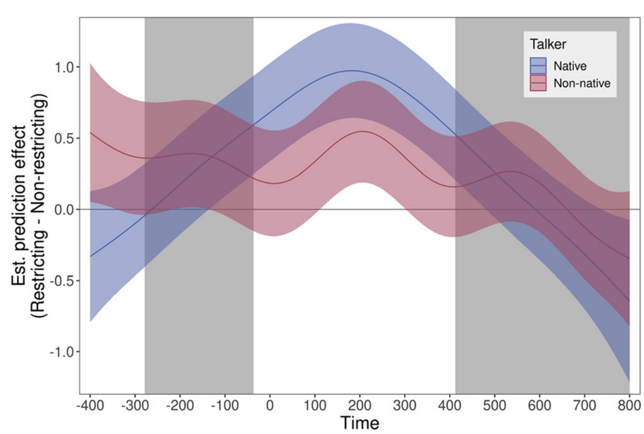

Fig. 2. Estimated prediction effect in response to the native and nonnative talkers over time, with 99% confidence intervals. The masks indicate time during which the difference between conditions was not statistically significant.

[7] We thank an anonymous reviewer for this observation.

Table 3. Summary of the generalized additive mixed model for prediction and experience

	Estimate	*SE*	*t* Value	*p* Value
Parametric coefficients				
Intercept	– 1.05	0.20	– 5.13	< .0001
Restricting	0.35	0.13	2.74	.0061
	Edf	Ref. *df*	*F* Value	*p* Value
Smooth terms				
Tensor: Time, Experience, NonRestricting	6.00	6.49	3.95	< .0001
Tensor: Time, Experience, Restricting	10.87	13.01	1.73	.0524
Smooth: Trial	1.00	1.00	0.70	.4026
Smooth: Frequency	1.00	1.00	0.01	.9281
Random effect: Time, Subject	365.24	538.00	45.19	< .0001
Random effect: Time, Item	145.92	214.00	34.94	< .0001
Random effect: Event	522.37	682.00	5.21	< .0001

The present results are consistent with both views, which are, in turn, consistent with the idea that prediction is automatic and requires effort. Further research will be necessary to clarify the exact nature of predictive processing during spoken language comprehension, and specifically at which levels such processing occurs, and when it begins.

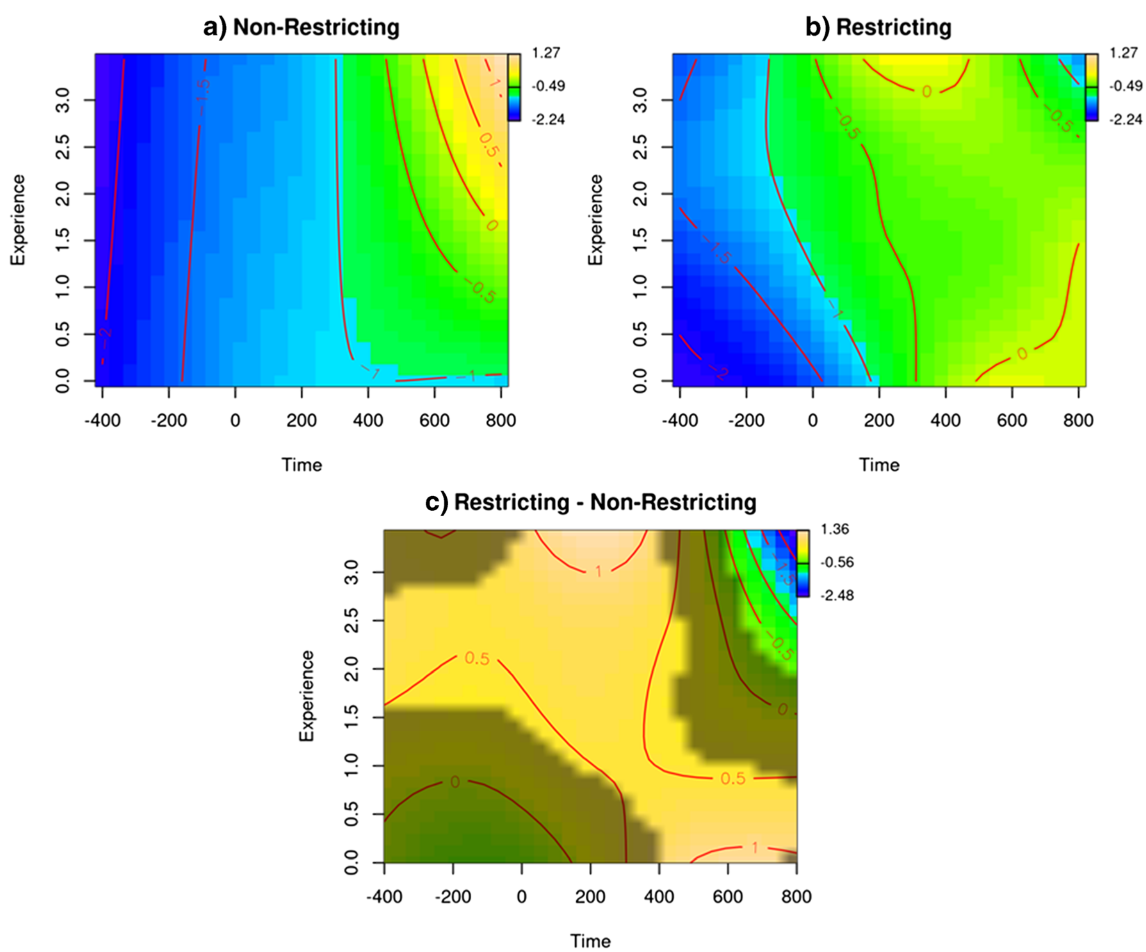

Fig. 3. Contour plots of time by accent experience for the nonrestricting condition (A) and the restricting condition (B). Dark gray (blue in the online color figure) indicates decreased looks to the target, whereas light gray (yellow in the online figure) indicates increased looks to the target. Panel C represents the difference; masked regions indicate areas that include zero within the 95% confidence interval, which are not significantly different from zero.

Analysis 2 demonstrated experience-dependent prediction, whereby greater lifelong experience with an accent resulted in a stronger prediction effect. Importantly, for the most experienced listeners, the pattern of prediction was visually similar to that for the native accent (cf. Figs. 2 and 3C). This suggests that accumulated experience with the variability associated with Chinese-accented English may ease the processing demands associated with signal decoding and lexical access, thus freeing up resources (or increasing certainty) for predictive processing. Reduced accent experience resulted in little to no prediction.

Huettig (2015) argued that mediating factors must be integrated into models of anticipatory language processing in order to comprehensively account for the data. Theories of predictive processing posit that predictions are based on prior experiences, at least at a global level. Verhagen, Mos, Backus, and Schilperoord (2018) showed that participants who differed in their usage-based experience of various registers differed in the expectations they generated for word sequences characteristic of each register. This suggests that participants have situational mental representations of language use that result in different predictions. Here, listener experience likely influences the certainty of the current input, which then affects (the engagement of) predictive processing. It is possible that a lack of control of individual accent experience has contributed to the inconsistency seen in studies examining N400 effects for foreign-accented speech (see the introduction). Self-reported accent experience, as assessed through a questionnaire (see the Appendix and Porretta et al., 2016), provides a quick and effective way to obtain estimates for investigating the effect of, or controlling for, prior exposure outside the laboratory.

It should be noted that, although the nonnative talker was very highly intelligible, he did differ in intelligibility from the native talker—as accentedness and intelligibility are known to covary (Porretta & Tucker, 2015). While prediction from nonnative speech is influenced by experience with the accent, this experience could also aid listeners in processing speech with lower intelligibility in general (native or nonnative). Further research will be required in order to clarify whether the effect of foreign-accented speech seen in the present data also occurs with native speech that varies in intelligibility.

In conclusion, this is the first demonstration that, when comprehending foreign-accented speech, listeners predict—albeit to a lesser extent than with native speech—prior to hearing a target word. Additionally, this prediction is enhanced by lifelong, accent-specific experience. Thus, predictive processing occurs even under suboptimal listening conditions, and individual differences in linguistic experience shape a listener's ability to predict during language processing.

Open Practices Statement None of the data or materials for the experiment reported here are publicly available; however, they can be made available to any qualified researcher upon request to the first author. The experiment was not preregistered.

Appendix

Table 4. Participant questionnaire

Question	Response Type
Gender	Male/female/other
Age	Free response
Have you been diagnosed with a hearing impairment?	Yes/no
Have you been diagnosed with a language disorder?	Yes/no
Are you native speaker of North American English?	Yes/no
Are you bilingual? (If yes, in which language(s)?)	Yes/no (Free response)
Which accent (or accents) did you hear during the experiment?	Free response
Have you ever studied Chinese?	Yes/no
Estimate your total lifetime experience interacting in English with speakers with a Chinese accent.	Percentage

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264. https://dx.doi.org/https://doi.org/10.1016/S0010-0277(99)00059-1

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H. (2010). The directed compound graph of English. An exploration of lexical connectivity and its processing consequences. In S. Olson (Ed.), *New impulses in word-formation* (Linguistische Berichte Sonderheft 17, pp. 383–402). Hamburg, Germany: Buske.

Baayen, R. H., van Rij, J., Cecile, D., & Wood, S. N. (2018). Autocorrelated errors in experimental data in the language sciences: Some solutions offered by Generalized Additive Mixed Models. In D. Speelman, K. Heylan, & D. Geeraerts (Eds.), *Mixed effects regression models in linguistics*. Berlin, Germany: Springer.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B.,… Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459. https://dx.doi.org/https://doi.org/10.3758/BF03193014

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474. https://dx.doi.org/https://doi.org/10.1016/j.jml.2007.09.002

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729. https://dx.doi.org/https://doi.org/10.1016/j.cognition.2007.04.005

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One*, *9*, e106953. https://dx.doi.org/https://doi.org/10.1371/journal.pone.0106953

Brouwer, S., & Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *Journal of Psycholinguistic Research*, *45*, 1151–1160. https://dx.doi.org/https://doi.org/10.1007/s10936-015-9396-9

Brown-Schmidt, S., & Toscano, J. C. (2017) Gradient acoustic information induces long-lasting referential uncertainty in short discourses. *Language, Cognition and Neuroscience*, *32*, 1211–1228. https://dx.doi.org/https://doi.org/10.1080/23273798.2017.1325508

Burchill, Z., Liu, L, & Jaeger, T. F. (2018). Maintaining information about speech input during accent adaptation. *PLoS One*, *13*, e0199358. https://dx.doi.org/https://doi.org/10.1371/journal.pone.0199358

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York, NY: Springer.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204. https://dx.doi.org/https://doi.org/10.1017/s0140525x12000477

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*, 92–102. https://dx.doi.org/https://doi.org/10.1016/j.bandl.2012.04.017

Henry, N., Hopp, H., & Jackson, C. N. (2017). Cue additivity and adaptivity in predictive processing. *Language, Cognition and Neuroscience*, *32*, 1229–1249. https://dx.doi.org/https://doi.org/10.1080/23273798.2017.1327080

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://dx.doi.org/https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*, 1–14. https://dx.doi.org/https://doi.org/10.1080/23273798.2015.1047459

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Lang & Ling Compass*, *2*, 647–670. https://dx.doi.org/https://doi.org/10.1111/j.1749-818x.2008.00072.x

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*, 32–59. https://dx.doi.org/https://doi.org/10.1080/23273798.2015.1102299

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, *60*, 661–671. https://dx.doi.org/https://doi.org/10.1080/17470210601183890

Porretta, V., & Kyröläinen, A.-J. (2019). Influencing the time and space of lexical competition: The effect of gradient foreign accentedness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 1832–1851. https://dx.doi.org/https://doi.org/10.1037/xlm0000674

Porretta, V., Kyröläinen, A.-J., van Rij, J., & Järvikivi, J. (2016). VWPre: Tools for preprocessing visual world data (Version 1.1.0, updated 2018-03-20). Retrieved from https://CRAN.R-project.org/package=VWPre

Porretta, V., & Tucker, B. V. (2015). Intelligibility of foreign-accented words: Acoustic distances and gradient foreign accentedness. In *Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 1–4). Glasgow, UK: University of Glasgow. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0657.pdf

Porretta, V., & Tucker, B. V. (2019). Eyes wide open: Pupillary response to a foreign accent varying in intelligibility. *Frontiers in Communication*, *4*, 8. https://dx.doi.org/https://doi.org/10.3389/fcomm.2019.00008

Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics*, *58*, 1–21. https://dx.doi.org/https://doi.org/10.1016/j.wocn.2016.05.006

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514–528. https://dx.doi.org/https://doi.org/10.1037/a0020990

Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, *85*, 245–255. https://dx.doi.org/https://doi.org/10.1016/j.neuropsychologia.2016.03.022

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215. https://dx.doi.org/https://doi.org/10.1037/0278-7393.6.2.174

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D.,… Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*, 247–250. https://dx.doi.org/https://doi.org/10.1016/j.jml.2004.03.002

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–553. https://dx.doi.org/https://doi.org/10.1037/0278-7393.19.3.528

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443–467. https://dx.doi.org/https://doi.org/10.1037/0278-7393.31.3.443

van Rij, J., Wieling, M., Baayen, R. H., & van Rijn, H. (2017). itsadug: Interpreting time series and autocorrelated data using GAMMs (R package version 2.2). Retrieved from https://CRAN.R-project.org/package=itsadug

Verhagen, V., Mos, M., Backus, A., & Schilperoord, J. (2018). Predictive language processing revealing usage-based variation. *Language and Cognition*, *10*, 329–373. https://dx.doi.org/https://doi.org/10.1017/langcog.2018.4

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

Wood, S. N. (2018). mgcv: Mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation (R package version 1.8-12). Retrieved from https://CRAN.R-project.org/package=mgcv

This research was supported by a Social Sciences and Humanities Research Council (http://www.sshrc-crsh.gc.ca/) Partnership grant (Words in the World, 895-2016-1008).