Bilingualism: Language and Cognition

cambridge.org/bil

Simultaneous French–German bilinguals' phonotactic knowledge and processing: Evidence from a well-formedness judgment and a lexical decision task

Leonardo Piot[1,2], Thierry Nazzi[2,3] and Natalie Boll-Avetisyan[1]

[1]Department of Linguistics, University of Potsdam, Germany; [2]Integrative Neuroscience and Cognition Center, Université Paris Cité, France and [3]Integrative Neuroscience and Cognition Center (INCC), CNRS, France

Research Article

Cite this article: Piot, L., Nazzi, T. and Boll-Avetisyan, N. (2026). Simultaneous French–German bilinguals' phonotactic knowledge and processing: Evidence from a well-formedness judgment and a lexical decision task. *Bilingualism: Language and Cognition* 1–16. https://doi.org/10.1017/S1366728926101667

Received: 29 May 2025 UTC
Revised: 27 May 2026 UTC
Accepted: 5 June 2026 UTC

Keywords:
phonotactics; simultaneous bilinguals; knowledge; processing; cross-linguistic

Corresponding author:
Leonardo Piot;
Email: leonardo.piot@uni-potsdam.de

T.N. and N.B.-A. share last authorship.

This research article was awarded Open Data badge for transparent practices. See the Data Availability Statement for details.

Abstract

Monolinguals have detailed knowledge of the phonotactic properties of their native language and apply such knowledge during speech processing. However, less is known about bilinguals' phonotactic knowledge and processing, notably in simultaneous bilinguals. Here, we investigated in two web-based experiments whether simultaneous French/German bilinguals show knowledge of their two languages' phonotactics (Experiment 1: well-formedness judgment task) and whether they apply such knowledge during speech processing (Experiment 2: lexical decision task). We compared the bilinguals to two groups of monolingually-raised French- and German-speaking adults. Analyses revealed that the bilinguals' phonotactic knowledge of their two languages is similar to that of the monolinguals, while they differ from the monolinguals in processing, showing reduced phonotactic sensitivity, possibly because of the co-activation of both languages' systems. These findings contribute to our understanding of how phonotactic properties are represented and used in simultaneous bilinguals' speech processing.

Highlights

- First study to explore simultaneous bilinguals' phonotactic knowledge and processing.
- Bilinguals show knowledge of their two L1s' phonotactics, similar to monolinguals.
- Bilinguals differ from monolinguals in phonotactic processing.

1. Introduction

Adults have detailed knowledge of the phonotactic regularities of their language(s), that is, the legal and probabilistic positioning and sequencing of speech sounds within the words of their language(s). This *phonotactic knowledge* is mainly assessed via well-formedness judgment tasks, where individuals are asked to rate the wordlikeness of nonwords varying in phonotactic probabilities and are thus probed to think about the properties of their language metalinguistically. Such studies generally show that monolinguals provide ratings that correlate with the phonotactic regularities in their languages: the more frequent the regularities in the nonwords, the better their wordlikeness (e.g., Bailey & Hahn, 2001; Frisch et al., 2001; Hay et al., 2004; Richtsmeier, 2011; Vitevitch et al., 1996). Besides this, phonotactic regularities have been shown to affect adult speech perception, facilitating speech processing. *Phonotactic processing* is studied with implicit tasks, designed to investigate the effect of phonotactics on online processing efficiency. Such studies have provided evidence that the frequency of segment sequences affects processing speed (Vitevitch & Luce, 1998, 2005). It also affects lexical access, as shown in lexical decision tasks in which listeners have to decide whether presented stimuli are real word in their language or not: monolingual speakers are quicker and more accurate to decide that stimuli are nonwords if they contain non-frequent phonotactic regularities (Vitevitch & Luce, 1999). Phonotactics also affects listeners' speech segmentation (Boll-Avetisyan & Kager, 2016; Dumay et al., 2002; McQueen, 1998; Mersad & Nazzi, 2011; van der Lugt, 2001), syllable processing (Smith & Pitt, 1999), and phoneme categorization (Massaro & Cohen, 1983; Pitt & McQueen, 1998). Importantly, studies testing listeners with unfamiliar languages have shown that phonotactic effects in speech processing, shown by native speakers, do not appear with naïve listeners (e.g., Otake et al., 1996). In fact, phonotactic effects can result in "perceptual repairs" when processing non-native speech (e.g., Dupoux et al., 2011; Hallé et al., 2008). For example, in languages in which it is phonotactically illegal, speakers perceptually repair the onset /tl/ into their legal onset /kl/ (e.g., French: Hallé et al., 1998; American-English: Hallé & Best, 2007). Thus, these effects are not primarily signal-driven; rather, they appear language-specific and experience-driven.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

The fact that phonotactics are acquired through experience raises the question of how bilinguals develop language-specific phonotactic knowledge and use it during processing. Indeed, bilinguals could be facing two challenges. First, they have to represent the phonotactic regularities of their two languages, even though some regularities might conflict between them (e.g., if a cluster is legal in one but illegal in other language). Second, they need to apply their knowledge dynamically as a function of language context, raising the question of how efficiently they activate/deactivate language-specific phonotactics during processing. These challenges might have particular impact on the representation and use of phonotactics in simultaneous bilinguals, who regularly use two languages they acquired from early infancy. In this study, we therefore aimed to investigate whether bilinguals know the phonotactics of both their languages (Experiment 1), and if so, whether they make equal use of such properties during speech processing (Experiment 2). To do so, we tested French–German simultaneous bilingual adults, and compared their behavior to monolingually-raised speakers of French and German.

Studies focussing on bilinguals' phonotactics have mainly investigated sequential bilinguals who have acquired one language from birth and a second one later in life[1] (during childhood: Anisfeld et al., 1969; Altenberg & Cairns, 1983; Carlson et al., 2016; Frisch & Brea-Spahn, 2010; Freeman et al., 2016, 2022; Sebastián-Gallés & Bosch, 2002; during adulthood: Boll-Avetisyan, 2011; Lentz & Kager, 2015; Trapman & Kager, 2009; Weber & Cutler, 2006). Studies investigating *phonotactic knowledge* suggest that bilingual adults possess knowledge of both their phonotactic systems and that they are able to use it appropriately based on the identity of the language processed (Altenberg & Cairns, 1983; Frisch & Brea-Spahn, 2010; but see Anisfeld et al., 1969). For example, Frisch and Brea-Spahn (2010) found that well-formedness judgments of spoken nonwords differing in phonotactic probability by early sequential Spanish-English bilingual adults did not differ significantly from those by monolingual speakers of English if nonwords were rated in English. Moreover, the bilinguals had nearly identical performance in wordlikeness ratings for both English and Spanish. This suggests that sequential bilinguals acquire both phonotactic systems at monolingual levels, and that such knowledge operates within each language independently during the judgment task.

Studies focusing on *phonotactic processing* suggest a more complex picture. On one hand, phonotactic effects are stronger when bilinguals process their dominant language, whether defined by greater language exposure in infancy/childhood, the language acquired first and/or language proficiency. For example, in a vowel detection gating task and an AX discrimination task, early sequential Spanish-English bilingual adults, like Spanish monolinguals, tend to repair word-initial sC clusters, which are unattested in Spanish, by adding a prosthetic /e/, but to a smaller extent and only when Spanish-dominant. Conversely, if they are English-dominant, they behave like English monolinguals and do not repair these clusters (Carlson et al., 2016). Moreover, Spanish-dominant early sequential Spanish-Catalan bilinguals show smaller sensitivity to Catalan phonotactics than Catalan-dominant bilinguals (Sebastián-Gallés & Bosch, 2002). Such an effect of proficiency was further documented in late learners of Dutch, as Dutch proficiency related to the size of the phonotactic effect in a Dutch lexical decision task (Lentz & Kager, 2015; Trapman & Kager, 2009) and in a short-term memory recall task (Boll-Avetisyan, 2011). On the other hand, sequential bilinguals show some interference of first language (L1) phonotactics during L2 processing, possibly because the L1 phonotactic system is not fully deactivated (Freeman et al., 2016, 2022; Lentz & Kager, 2015; Weber & Cutler, 2006). For example, Weber and Cutler (2006) showed that late sequential bilinguals, although they have acquired their L2 phonotactics, are not able to fully deactivate their L1 phonotactics. This interference effect of the L1 during L2 processing might also be related to the dominance effect described earlier: not only is phonotactic processing stronger in the dominant language, but perhaps it is harder for bilinguals to activate only L2 phonotactics during processing. An alternative explanation for these interference effects is that phonotactic processing is *language-nonselective*, which means that the bilinguals' two phonotactic systems are always activated in parallel during processing. The latter explanation aligns with what is known about bilinguals' lexical processing, which is also largely language-nonselective (e.g., De Groot & Nas, 1991; Marian & Spivey, 2003; Van Assche et al., 2020).

Only one study tested both *phonotactic knowledge* and *processing* in sequential bilinguals. Altenberg and Cairns (1983) investigated whether the phonotactic systems of German–English bilinguals, most of whom learned both languages before age 12 but were overall more proficient in English, interact during language processing. To test this, they used two tasks: a well-formedness judgment task and a lexical decision task, and presented participants with written nonwords that were phonotactically legal in one language while illegal in the other. Bilinguals were tested either in an English or in a German mode (i.e., the language to be considered for the task was either English or German, as were the instructions), and their performance in the English mode was compared with that of English monolinguals. In the judgment task, bilinguals showed language-specific adaptation: when tested in English, there was no evidence that their responses differed from those of English monolinguals. The authors interpreted this null result as suggesting that bilinguals' phonotactic knowledge was correctly applied to the specific language context, with no cross-language interference. However, the lexical decision task revealed cross-language influence in phonotactic processing: bilinguals tested in English were slower than monolinguals to reject nonwords that were phonotactically legal in German but illegal in English. These results suggest that bilinguals' phonotactic systems are not fully independent but rather remain jointly activated and can interact during processing: even the non-dominant language interfered with processing in the dominant language.

In summary, the above studies highlight two important aspects of sequential bilinguals' phonotactics. First, sequential bilinguals' experience does not appear to significantly influence their *phonotactic knowledge* in their two languages, as assessed with a metalinguistic judgment task, suggesting that their phonotactic representations are detailed and flexible enough to be selectively accessed to the appropriate language context. Second, the two *phonotactic processing systems* of sequential bilinguals do not seem to be completely independent from one another, as revealed by tasks requiring online processing and showing the influence of one language-specific system on the other. In addition to co-activation of the two phonotactic systems, sequential bilinguals also show evidence of language-dominance effects: the more proficient they

[1]Throughout the manuscript, following Sebastián-Gallés, Bosch, and colleagues (e.g., Sebastián-Gallés & Bosch, 2002), we use the term *simultaneous bilinguals* to refer to bilinguals who were exposed to both of their languages from birth or before the age of 2 years, *early sequential bilinguals* to refer to the bilinguals who learned their L2 during childhood but were exposed to one language only during minimally their first few years of life, and *late sequential bilinguals* to refer to the bilinguals who learned their L2 in adolescence or adulthood.

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

are with one language, the stronger the phonotactic effects in processing that language.

Would the findings on phonotactic knowledge and processing extend to simultaneous bilinguals who were exposed to their two native languages from early infancy? Or could it be that more experience with the two languages, especially during a developmental period critical for phonological acquisition, helps better navigate between the two phonotactic systems during processing and mitigates dominance effects? To our knowledge, only one study focusing on phonotactic processing investigated simultaneous bilingual adults, comparing them to early and late sequential bilinguals, as well as monolingual controls (Parlato-Oliveira et al., 2010). This study investigated the identity of the epenthetic vowel repairing illegal consonant clusters in different groups of Japanese/Brazilian-Portuguese bilinguals (i.e., /u/ is the typical epenthetic vowel used by Japanese listeners, /i/ the typical vowel used by Brazilian listeners). It did not find clear differences between simultaneous and sequential bilinguals, who all patterned with the Brazilian monolinguals. The authors interpret these findings as resulting from bilinguals' current linguistic proficiency and exposure, rather than language experience in infancy, given that all bilingual groups were Brazilian-dominant and living in Brazil. These findings suggest that current language exposure and use, rather than language input in childhood, influence phonotactic processing. Nevertheless, more studies are needed to better understand whether simultaneous bilinguals know the phonotactics of their two native languages and show interference of one language phonotactics when processing the other language. These aspects will be investigated in the current study.

1.1. The current study

In the present study, we used a similar design as Altenberg and Cairns (1983) to investigate whether the dichotomy found in sequential bilinguals applies to simultaneous bilinguals, that is, *phonotactic knowledge* of their two languages being similar to monolinguals' knowledge (Experiment 1), and *phonotactic processing* showing interference between their two languages (Experiment 2). Accordingly, bilinguals' *phonotactic knowledge* was assessed through well-formedness judgments and *phonotactic processing* through lexical decisions, allowing to distinguish across tasks between phonotactic knowledge and online processing effects. The study's materials were also inspired by their study, such that we manipulated the phonotactic properties of our stimuli in four ways: nonwords beginning with clusters that were (1) frequent or (2) non-frequent in both French and German, (3) frequent in French but non-frequent in German, or (4) frequent in German but non-frequent in French. This allowed us to assess whether phonotactic well-formedness in one language influences bilinguals' behavior when perceiving the other language.

Besides our focus on simultaneous rather than sequential bilinguals, our study differed from Altenberg & Cairns' study as follows. First, we presented auditory – as opposed to written – stimuli, given that a study found differences in phonotactic effects between the two modalities (Bailey & Hahn, 2001). Second, we used stimuli pronounced by both a native speaker of French and a native speaker of German, to test whether pronunciation modulates the degree of cross-linguistic phonotactic activation in speech processing, given studies showing that language-specific pronunciations modulate bilinguals' degree of parallel activation of their lexicons (Schulpen, Dijkstra, Schriefers & Hasper, 2003; Lagrou, Hartsuiker, & Duyck, 2013; but see Lagrou, Hartsuiker, & Duyck, 2011). Third, we tested two monolingual-raised control groups, to validate our experiments and to directly test how bilinguals' phonotactic effects compared to monolingual-raised speakers of their two languages. We initially expected that simultaneous bilinguals would show an effect of interference between their two languages in both drawing on knowledge (affecting their well-formedness judgments) and in using phonotactics for processing (affecting their lexical decisions). We are aware that for *phonotactic knowledge*, this hypothesis, which we included in our pre-registration, goes against the findings on sequential bilinguals reviewed above (Altenberg & Cairns, 1983; Frisch & Brea-Spahn, 2010), and might be difficult to test with a metalinguistic judgment task capturing conscious knowledge rather than automatic processing, which limits its interpretative value regarding co-activation-based interference. However, we reasoned that it would be important to assess whether perhaps, with changes in design and methodology (the inclusion of simultaneous bilinguals and two monolingually-raised control groups, spoken stimuli and two pronunciations), we would find evidence for subtle interference between the bilinguals' two L1s. Nevertheless, given the clearer evidence for phonotactic interference in *processing* than in drawing on *knowledge* found in sequential bilinguals, we expected that if task-specific effects arose, they would go in the same direction (i.e., interference would emerge in the lexical decision task). In terms of pronunciation cues, we hypothesized that if pronunciation cues play a role in the degree of cross-linguistic phonotactic activation, then the interference effect might be more likely to occur when language-specific pronunciations would not match the language being processed. These hypotheses were pre-registered on osf (well-formedness judgment task: https://osf.io/yda94; lexical decision task: https://osf.io/aep28) and included power analyses to justify our sample size for the bilingual group of participants, as well as the statistical analysis plans.

2. Experiment 1: Well-formedness judgment task

2.1. Methods

2.1.1. Participants

Thirty-one simultaneous French–German bilinguals (mean age = 26.39 years, range = 18–40 years, 21 females, 10 males), 30 French monolingually-raised (mean age = 29.53 years, range = 20–40 years, 25 females, five males), and 35 German monolingually-raised participants (mean age = 22.82 years, range = 18–34 years, 28 females, seven males) were tested and included in the analysis. None had hearing, language, or cognitive impairments, and all provided correct answers during the attention checks (see procedure below). Bilinguals had to be exposed to both French and German regularly before 2 years of age. Monolingually-raised participants were defined as individuals exclusively exposed to one language (French or German) from birth to at least 6 years of age. We did not collect information about languages learned later in life in the monolingually-raised groups, although formal foreign-language instruction is widespread in French and German schools. Thus, the term "monolingually-raised" refers to early monolingual language experience rather than the absence of later foreign-language knowledge. One additional bilingual was tested but excluded from the analyses due to experiment error. Eleven additional monolingually-raised participants were excluded from the analyses because they provided incorrect answers during the attention checks. A descriptive table characterizing the bilingual sample in language experience and proficiency can be found in Supplementary Appendix S2 (Table S3).

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

2.1.2. Materials

We created 64 nonwords divided into four sets differing in phonotactic properties. Two sets had convergent phonotactic properties across French and German, and two had divergent phonotactic properties between the two languages. Each nonword had a #**CC**VCVC# structure and had no direct lexical neighbor in French or German. Phonotactic differences between sets depended on the word-initial consonant cluster: in the *FreqBoth* set, the four word-initial clusters (/tr/, /dr/, /fl/, and /bl/) were frequent in both French and German. In the *FreqNative*/*InfreqNative* sets, the word-initial clusters were frequent in one language, but non-frequent[2] in the other. Hence, a stimulus in the *FreqNative* set for French was also in the *InfreqNative* set for German and vice versa, such that nonword status depended on the target language (see procedure below for more details). The four clusters phonotactically frequent in French while non-frequent in German were /st/, /sp/, /dj/, and /vr/. Conversely, the clusters that are frequent in German but non-frequent in French were /kn/, /kv/, /ʃp/, and /ʃt/. Finally, in the *InfreqBoth* set, word-initial clusters were non-frequent in both French and German: /ʃk/, /sr/, /sb/, and /fn/. Thus, four different word-initial clusters were used for each phonotactic set (16 in total). Each cluster was combined with four out of eight possible tails (i.e., /agik/, /opik/, /iton/, /orɛt/, /eʃɛt/, /usap/, /atyl/, /abɛt/; see Supplementary Appendix S1 for details about the stimuli) to create our final list of stimuli. Each tail was distributed evenly across the four phonotactic sets, ensuring that any difference between sets could only be explained by differences in phonotactic frequency of the word-initial clusters.

One female native speaker of French and one of German recorded all stimuli in the same soundproof booth using the same equipment. We normalized the stimuli's mean intensity at 70 dB using PRAAT (Boersma, 2001). The average duration and pitch of nonwords for each phonotactic set and speaker are reported in Table 1. The French stimuli did not have lexical stress, which does not exist in French, whereas the German stimuli followed a strong-weak stress pattern, typical for German.

2.1.3. Procedure, apparatus, and design

The experiment was conducted online using *Labvanced* (Finger et al., 2017).

Table 1. Average duration and pitch of the nonwords, divided by pronunciation and phonotactic group, as well as the real words used in Experiment 2

	French pronunciation		German pronunciation	
	Duration (s)	Pitch (Hz)	Duration (s)	Pitch (Hz)
FreqBoth	0.629	241	0.644	253
FR: FreqNative DE: InfreqNative	0.666	238	0.750	266
DE: FreqNative FR: InfreqNative	0.658	224	0.740	278
InfreqBoth	0.680	239	0.762	273
German words	0.599	214	0.700	226
French words	0.641	230	0.778	231

[2]We use the term "non-frequent" for clusters that are illicit or strongly marked but nevertheless could occur in some loanwords/borrowings. We expect these patterns to be among the most strongly disfavored by the listeners.

Monolingual versions: Monolingually-raised participants were instructed that on each trial, they would hear sounds and were asked to rate how acceptable these sounds would be as a new word in French/German, using a scale from 1 ("Not acceptable at all") to 7 ("Highly acceptable"). The experiment started with eight training trials and was followed by 128 test trials (i.e., the 64 nonwords in the two different pronunciations). A trial ended once the participant responded, or after 6 seconds. Because the task was web-based, four attention checks required participants to type the stimulus from the preceding trial. In these checks, responses were classified as incorrect when they indicated clear inattention (e.g., omitted responses or responses unrelated to the target item). Minor phonemic deviations (e.g., vowel substitutions or consonant changes within clusters) were not treated as errors, as such responses were taken to reflect successful perception of the stimulus despite small transcription differences. Pronunciation was blocked and presented in a fixed order, such that in the first half of the task (i.e., the first block) we presented the stimuli pronounced by the native speaker of the target language (*target pronunciation*) and in the second half of the task (i.e., the second block) we presented the same stimuli pronounced by the other speaker (*non-target pronunciation*). Trial order was pseudorandomized to avoid presenting more than two consecutive nonwords with the same word-initial cluster or more than four nonwords with the same tail in a sequence.

Bilingual version: Bilinguals completed this task in their two L1s. At the beginning, the procedure was exactly the same as for monolingually-raised participants: they completed the training and experimental trials and were asked to judge the acceptability of nonwords as new words of a given language, getting the instructions in that language. Once the participants had completed the 128 trials in that language, a written message stating that half of the task was completed was presented in both French and German. At this point, participants could take a break before resuming the experiment by pressing a key. Once resumed, the instructions appeared in the other language and prompted participants to judge the acceptability of nonwords as new words of that second language. We counterbalanced the order of target language across participants, such that half started with target language French while the other half started with target language German.

At the end of the experiment for the monolingually-raised participants, and after each language block for the bilinguals, participants completed the vocabulary questionnaire Lextale (German version: Lemhöfer & Broersma, 2012; French version: Brysbaert, 2013) assessing their proficiency in the tested language. Lextale consists of a list of written items (i.e., words and pseudowords) for which participants indicated whether they were existing words in the language. Finally, bilingual but not monolingually-raised participants were interviewed by phone on their language background (using a modified version of the LEAP-Q: Marian et al., 2007, mixed with the questionnaire used in Dupoux et al., 2010, as in Boll-Avetisyan et al., 2020) a few days after the experiment.

2.1.4. Predictions and data analysis

We predicted higher ratings for *FreqBoth* than *InfreqBoth*, and *FreqNat* than *InfreqNat*, such that the more frequent the phonotactic properties of the items, the higher their wordlikeness. Crucially, because we predicted phonotactic knowledge for both the monolingually-raised participants in their respective native languages, and the bilinguals in each of their languages, we predicted that both monolingually-raised participants and bilinguals would rate higher the frequent phonotactic conditions compared to the non-frequent conditions. Even though well-formedness judgments capture explicit, metalinguistic knowledge rather than automatic online processing (and are thus not a direct

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

test of language co-activation), we nevertheless predicted that bilingual and monolingually-raised participants could differ in the size of phonotactic effects across conditions (notably between *FreqNat* and *InfreqNat*, where the phonotactic properties of their two languages diverged), and that language-specific pronunciation might differently affect monolingually-raised and bilingual listeners' phonotactic well-formedness intuitions, attesting subtle differences in phonotactic knowledge.

To test those predictions, we analyzed participants' rating in each trial with a (logit) ordinal mixed-effects model using the *ordinal* package (Christensen, 2019) in R (R Core Team, 2021; see Liddell & Kruschke, 2018 and the supplementary materials in Oh et al. 2020, for more details about this kind of analysis). Thus, the dependent variable was the participants' rating on a Likert scale (1–7). The fixed part of the model included the main effects of three factors and their interactions: phonotactics (four levels, coded with sliding contrasts that compared *InfreqBoth* with *FreqBoth*, *FreqBoth* with *FreqNative*, and *FreqNative* with *InfreqNative*), pronunciation (two levels, sum-contrast coded, with an item being pronounced by the *native* versus *non-native* speaker of the target language, the target language being the one to which participants had to relate their wordlikeness ratings), and language status (two levels, *monolingual* versus *bilingual*, sum-contrast coded). The random structure was kept as maximal as possible (Barr et al., 2013) and included by-participants random slopes for phonotactics in interaction with pronunciation, as well as by-items random slopes for pronunciation. The resulting formula was as follows:

$$\begin{aligned} Rating \sim\ & Phonotactics{*}Language\ Status{*}Pronunciation \\ & + (1 + Phonotactics{*}Pronunciation|Participant) \\ & + (1 + Pronunciation|Item). \end{aligned}$$

Post-hoc comparisons of estimated marginal means (EMMs) were computed with the *emmeans* package (Lenth, 2024) and visualizations with the *ggplot* package (Wickham, 2016) in R.

Deviation from the pre-registration: When pre-registering the lexical decision task (Experiment 2), we realized that the pre-registered statistical model for Experiment 1 was not ideal. Our pre-registered model tested whether phonotactic properties, language-specific pronunciation, language status and their interaction affected participants' ratings for each target language separately (French or German) with the following R-formula:

$$\begin{aligned} Rating \sim\ & TargetLang/(Phonotactics{*}Pronunciation{*} \\ & Language\ Status) + (1 + Phonotactics{*}Pronunciation| \\ & Participant) + (1 + Pronunciation|Item). \end{aligned}$$

We provide the results of this model in Supplementary Appendix S3 (Table S7). However, considering our central hypotheses, this model was unnecessarily complex and difficult to interpret, as we were interested in the effect across the two languages. Thus, the present analyses follow the model pre-registered for Experiment 2. Furthermore, because our phonotactic factor has four levels, we initially planned to test the differences between *FreqBoth-FreqNat*, *FreqNat-InfreqNat*, and *InfreqNat-InfreqBoth*, expecting bilingualism effects only between *FreqNat-InfreqNat*. We later realized that testing differences between both *FreqBoth-InfreqBoth* and *FreqNat-InfreqNat* provided a more informative and concise analysis allowing more robust interpretations. Therefore, we used these contrasts here (and in Experiment 2, for which they were pre-registered).

2.2. Results

The results of the statistical model are presented in Table 2; a plot summary of the raw data is presented in Figure 1. These reveal that there were statistically significant main effects of the *FreqBoth-InfreqBoth* ($p < .001$) and *InfreqNat-FreqNat* ($p < .001$) contrasts: in both contrasts, participants rated the nonwords containing frequent clusters as more wordlike than nonwords containing non-frequent ones. There was also a main effect of pronunciation ($p < .001$), with participants rating the nonwords as more wordlike when pronounced by the native speaker of the target language. There was an interaction of phonotactics and pronunciation in both the *FreqBoth-InfreqBoth* ($p < .001$) and *InfreqNat-FreqNat* ($p < .05$) contrasts, indicating that the phonotactic effect was modulated by the native/non-native status of the speaker pronouncing the stimuli, with a larger effect if the nonwords were pronounced by the native speaker of the target language compared to when they sounded non-native. Finally, there was a significant interaction between the *FreqBoth-InfreqBoth* contrast and language status ($p < .001$), indicating that the phonotactic effect was smaller in bilinguals than monolingually-raised participants. Follow-up post-hoc tests per language group confirmed that both groups reliably distinguished the phonotactic conditions (the full output of orthogonal pairwise comparisons can be found in Supplementary Table S4). For monolingually-raised participants, *FreqBoth* was rated higher than *InfreqBoth* (log-odds = 1.45, OR = 4.26, $p < .001$) and *FreqNat* was rated higher than *InfreqNat* (log-odds = 1.12, OR = 3.05, $p < .001$). The same pattern was observed in the bilingual group, although with smaller effect sizes: *FreqBoth* was rated higher than *InfreqBoth* (log-odds = 0.78, OR = 2.19, $p < .001$), and *FreqNat* was rated higher than *InfreqNat* (log-odds = 0.87, OR = 2.38, $p < .001$). Thus, the effects of phonotactics were statistically significant in both monolingually-raised and bilingual participants, but the magnitude of the phonotactic effect also differed across groups, establishing subtle differences in phonotactic knowledge. No other main effects or interactions reached significance.

2.3. Discussion

Experiment 1 assessed simultaneous French–German bilinguals' phonotactic knowledge with a well-formedness judgment task and compared it to that of monolingually-raised participants. Overall, our findings establish that, as predicted, all groups of participants were sensitive to the phonotactic properties of their native languages: the more frequent the phonotactic regularities in the nonwords, the better their wordlikeness. This is consistent with previous findings showing effects of language-specific phonotactic properties on well-formedness judgments by monolingual adults (e.g., Vitevitch et al., 1996) and sequential bilinguals (Altenberg & Cairns, 1983; Frisch & Brea-Spahn, 2010). Additionally, it shows that phonotactic differences restricted to word-initial clusters, as implemented here, are sufficient to elicit such effects. It also provides evidence that the task materials and procedure were valid and allowed the investigation of phonotactic knowledge in adults raised in monolingual or bilingual environments.

Turning to group comparisons, the results indicate that simultaneous bilinguals distinguished between the frequent and non-frequent phonotactic conditions in our task, as did monolingually-raised speakers. Only a small group difference emerged: a smaller effect in the bilingual compared to the monolingually-raised group in the *InfreqBoth-FreqBoth* contrast. While the source of this difference is unclear and might signal slight differences in (drawing on) phonotactic knowledge, it should be interpreted in light of the overall findings that

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

Table 2. Ordinal mixed-effects model of well-formedness ratings, taking into account the phonotactic conditions, language-specific pronunciation, and language status

	Parameter	Estimate	Std. Error	*z*	*P*
Effects	FreqBoth-Infreqboth	1.115	0.145	7.692	<.001***
	FreqNat-Freqboth	−0.103	0.117	−0.874	.382
	InfreqNat-FreqNat	−0.991	0.084	−11.839	<.001***
	Pronunciation	1.873	0.201	9.335	<.001***
	LangStatus	0.266	0.274	0.972	.331
	FreqBoth-Infreqboth × pronunciation	0.576	0.160	3.597	<.001***
	FreqNat-Freqboth × pronunciation	−0.143	0.139	−1.025	.305
	InfreqNat-FreqNat × pronunciation	−0.212	0.099	−2.143	.032*
	FreqBoth-Infreqboth × LangStatus	−0.667	0.167	−3.999	<.001***
	FreqNat-Freqboth × LangStatus	0.201	0.114	1.760	.078
	InfreqNat-FreqNat × LangStatus	0.248	0.164	1.517	.129
	pronunciation × LangStatus	−0.023	0.391	−0.058	.954
	FreqBoth-Infreqboth × pronunciation × LangStatus	−0.099	0.192	−0.517	.605
	FreqNat-Freqboth × pronunciation × LangStatus	−0.106	0.170	−0.622	.534
	InfreqNat-FreqNat × pronunciation × LangStatus	0.117	0.187	0.625	.532
Thresholds	1|2	−1.820	0.145		
	2|3	−0.183	0.139		
	3|4	0.815	0.139		
	4|5	1.468	0.144		
	5|6	2.517	0.145		
	6|7	4.030	0.149		

Note: The two-level factors were coded with sum contrasts (i.e., pronunciation: non-target_pronun = −0.5/target_pronun = 0.5; LangStatus: Monolingual = −0.5/Bilingual = 0.5), while the factor phonotactic was coded with sliding contrasts (i.e., *InfreqBoth. vs. FreqBoth, FreqBoth vs. FreqNative, and FreqNative vs. InfreqNative*). *: $p \leq .05$, **: $p \leq .01$, ***: $p \leq .001$.

bilinguals showed robust sensitivity to phonotactic patterns in both of their languages. These findings extend those found with sequential bilinguals, showing that their phonotactic knowledge is not significantly different from monolinguals (Altenberg & Cairns, 1983; Frisch & Brea-Spahn, 2010). Thus, the hypothesis that co-activation would affect bilinguals' phonotactic well-formedness intuitions was not supported by our data. We acknowledge that this was a difficult hypothesis to probe, given that the task was designed to capture conscious knowledge rather than automatic processing, limiting its interpretative value regarding co-activation-based interference.

Furthermore, we found that pronunciation cues modulated phonotactic well-formedness intuitions in both groups (smaller effects when items were pronounced by a speaker of the non-target language). However, we found no evidence that the language-specific pronunciation impacted bilinguals and monolingually-raised individuals differently. This suggests that pronunciation effects likely reflect general processing demands, with reduced effects under more demanding/unexpected pronunciations. We return to this effect in the General Discussion.

In sum, while Experiment 1 used a metalinguistic judgment task to establish that simultaneous bilinguals possess detailed phonotactic knowledge in both languages, Experiment 2 used a lexical decision task to investigate in the same bilinguals whether such knowledge is used during online phonotactic processing, and whether bilinguals show evidence of co-activation of their two phonotactic systems during processing.

3. Experiment 2: Lexical decision task

3.1. Methods

3.1.1. Participants

Twenty-three bilinguals (mean age = 25.7 years, range = 19-34 years, 13 females, 10 males) participated in the lexical decision task and were included in the analyses. They had participated in Experiment 1 at least 1 year before participating in Experiment 2. Additionally, 27 French monolingually-raised (mean age = 28.8 years, range = 18–40, 19 females, eight males) and 31 German monolingually-raised (mean age = 25.77, range = 18–40 years, 16 females, eight males, eight other/prefer not to say) participants who had not previously participated in Experiment 1 were tested and included in the analyses. Participants' inclusion criteria were as in Experiment 1, except for the attention checks, which were not implemented in Experiment 2.

3.1.2. Materials

The nonword stimuli were a subset of the stimuli of Experiment 1[3] (32 out of 64 nonwords, see Supplementary Appendix S1 for more details). We also used 32 real German and 32 real French words, starting with some of the same word-initial clusters as the

[3]An analysis of the data of Experiment 1 restricted to these items showed results that were very similar to the main confirmatory model, in terms of both effect sizes and *p*-values.

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

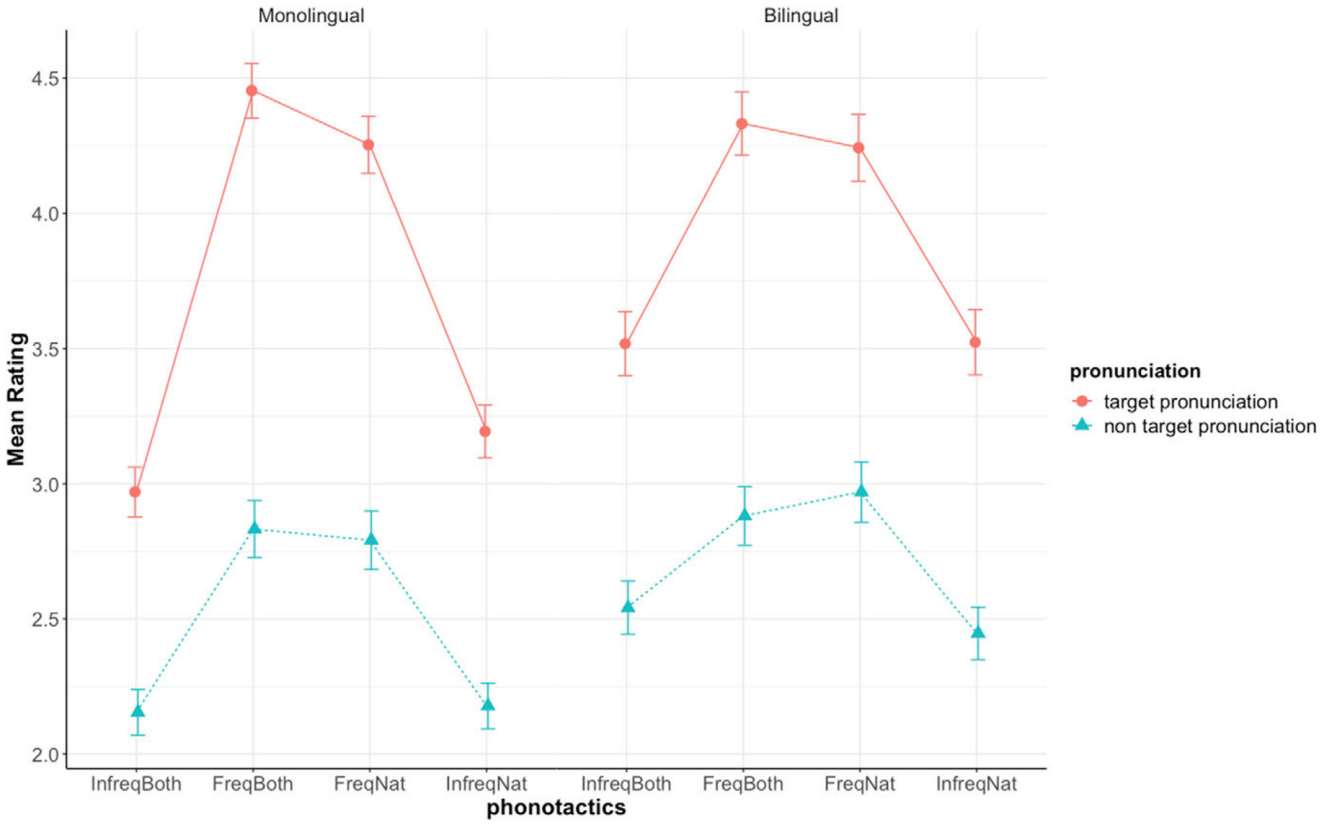


Figure 1. Mean well-formedness ratings (and CIs) in relation to phonotactic condition (*InfreqBoth*/*FreqBoth*: clusters non-frequent/frequent in both French and German, respectively; *FreqNat*/*InfreqNat*: clusters frequent/non-frequent in the target language and non-frequent/frequent in the other language, respectively) split by language-specific pronunciation (target pronunciation, non-target pronunciation) and language status (monolingual, bilingual).

nonwords, as fillers. The words were pronounced by the same speakers who pronounced the nonwords, and audio files were processed as in Experiment 1. The average duration and pitch of the German and French words spoken by the French and the German native speaker can be found in Table 1.

3.1.3. Procedure, apparatus, and design

The experiment was conducted online using either *Labvanced* (Finger et al., 2017) for the monolingual French version or *lab.js* (Henninger et al., 2020) for the monolingual German and the bilingual versions. After the lexical decision task, monolingually-raised participants completed the language-appropriate LexTale vocabulary questionnaire (Brysbaert, 2013; Lemhöfer & Broersma, 2012). Bilinguals did not complete it again, having already done so for Experiment 1.

Monolingual versions: Participants were instructed that on each trial, they would hear a word and were asked to decide as quickly and accurately as possible whether what they heard was a real word or not by pressing a key. The task started with eight practice trials with feedback ("correct!"/"incorrect!" on screen). Following this, the test phase consisted of two blocks of 64 trials, half presenting real words (in German or French, depending on the target language) and half presenting nonwords. Within each block, trial order was pseudorandomized to avoid presenting more than two consecutive trials of the same lexical (i.e., word/nonword), phonotactic, or pronunciation condition.

Bilingual version: Bilinguals completed this task in their two L1s. At the beginning, the procedure was exactly the same as for monolingually-raised participants: after the practice phase, they were tested in one language, receiving instructions in that language. Once they had completed the 128 trials in that language, a written message stating that half of the task was completed was presented in both French and German. At this point, participants could take a break before resuming the experiment by pressing a key. Once resumed, they were tested and received instructions in the other language. We counterbalanced the order of target language across participants, such that half started with French while the other half started with German.

3.1.4. Predictions, data preprocessing, and analysis

We predicted that accuracy in rejecting a nonword as being a real word in the language would be higher for *InfreqBoth* than *FreqBoth*, and *InfreqNat* than *FreqNat*. We also predicted a significant interaction between language status and phonotactics, or between language status, phonotactics, and pronunciation, resulting from co-activation of bilinguals' two phonotactic systems.

As pre-registered, before analyzing the data, we excluded trials during which participants responded too quickly (<100 ms; Vitevitch & Luce, 1999) or too slowly (>2500 ms; Lentz & Kager, 2015). Based on this criterion, 131 trials out of 6656 trials (1.97%) were excluded from the analysis. Furthermore, we excluded trials presenting real words, since those were only used as fillers in the task.

Our pre-registered analysis focused on the accuracy data of the nonwords with a (logit) generalized mixed-effects model. The independent factors and the model structure were the same as in Experiment 1. The random effects were grouped by participant and item. For this, we used R with the packages *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) to obtain *p*-values. Because the model did not converge when including random slopes, we only

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

included random intercepts for participants and items. The formula was as follows:

$$Accuracy \sim Phonotactics{*}Language\ Status{*}Pronunciation + (1|Participant) + (1|Item).$$

As suggested by a reviewer, in addition to the pre-registered accuracy model, we analyzed participants' reaction times (RTs) in trials where they accurately rejected the nonwords (excluding 373–5.72% – inaccurate trials). Because the duration of the auditory stimuli was variable, RTs were corrected on a trial-by-trial basis by subtracting the duration of the spoken nonword from the raw response latency (e.g., McQueen, 1998). This correction provides estimates of response time relative to stimulus offset, thereby reducing variance attributable to differences in item durations. All subsequent analyses were conducted on these corrected RTs. Then, the RTs were screened for outliers using a model-based procedure that accounts for the hierarchical structure of the data. Specifically, we computed an intercept-only linear mixed-effects model including the full random-effects structure, with random intercepts and slopes for participants and items. This model captures subject/item-level variability in baseline response speed without accounting for experimental effects. Scaled residuals from this model were then computed, and observations with absolute residuals exceeding 2.5 standard deviations were excluded from further analyses (a total of 205 trials were excluded −3.33% – for a similar procedure, see Baayen & Milin, 2010). The proportion of excluded data was comparable across phonotactic conditions.

We used linear mixed-effects models to analyze RTs. The model structure was the same as for the accuracy model, except that random slopes for participants and items were included. The formula was as follows:

$$CorrectedRT \sim Phonotactics{*}Language\ Status{*}Pronunciation + (1 + Phonotactics|thickmathspaceParticipant) + (1 + Pronunciation|thickmathspaceItem).$$

3.2. Results

3.2.1. Accuracy analyses

The results are presented in Table 3 and a plot summary of the raw data in Figure 2. They establish a statistically significant main effect of *FreqBoth-InfreqBoth* ($p = .041$) and *InfreqNat-FreqNat* ($p < .001$) indicating that participants were more accurate to decide that a stimulus was not a word in their language in the *InfreqBoth* than *FreqBoth* conditions, and the *InfreqNat* than *FreqNat* conditions. While the main effect of language status was not significant ($p = .647$), it significantly interacted with phonotactics (*FreqBoth-InfreqBoth*: $p = .003$; *InfreqNat-FreqNat*: $p = .002$), indicating that phonotactic properties modulated participants' accuracy differently between the monolingually-raised and bilingual groups. Post-hoc comparisons (the full output of orthogonal pairwise comparisons can be found in Supplementary Table S5) revealed that monolingually-raised participants' accuracy significantly differed between *FreqBoth* (EMM = 0.903, 95%CI [0.837; 0.945]) and *InfreqBoth* (EMM = 0.971, 95%CI [0.943; 0.985]; z = 3.117, $p = .010$), and between *FreqNat* (EMM = 0.920, 95%CI [0.872; 0.951]) and *InfreqNat* (EMM = 0.983, 95%CI [0.967; 0.991]; z = −5.872, $p < .001$). There were no significant differences between these conditions in the bilingual participants (*FreqBoth*: EMM = 0.958, 95%CI [0.913; 0.980], *InfreqBoth*: EMM = 0.968, 95%CI [0.932; 0.985], z = 0.656, $p = 0.914$; *FreqNat*: EMM = 0.947, 95%CI [0.901; 0.973], *InfreqNat*: EMM = 0.967, 95%CI [0.935; 0.984], z = −2.041, $p = 0.173$). Finally, there was a significant

Table 3. Generalized mixed-effects model of decision accuracy for nonwords, taking into account the phonotactic conditions, language-specific pronunciation, and language status

	Parameter	Estimate	Std. Error	*z*	*p*
Effects	(Intercept)	3.766	0.213	17.689	< .001***
	FreqBoth – InfreqBoth	−0.793	0.389	−2.039	.041*
	FreqNat – FreqBoth	−0.016	0.334	−0.047	.962
	InfreqNat – FreqNat	1.061	0.185	5.744	<.001***
	Pronunciation	0.139	0.126	1.104	.270
	LangStatus	0.152	0.333	0.457	.647
	FreqBoth – InfreqBoth × pronun	−0.099	0.344	−0.288	.773
	FreqNat – FreqBoth × pronun	−0.700	0.295	−2.378	.017*
	InfreqNat – FreqNat × pronun	0.054	0.368	0.148	.883
	FreqBoth – InfreqBoth × LangStatus	1.019	0.344	2.965	.003**
	FreqNat – FreqBoth × LangStatus	−0.464	0.295	−1.574	.116
	InfreqNat – FreqNat × LangStatus	−1.122	0.369	−3.040	.002**
	pronun × LangStatus	0.346	0.252	1.373	.170
	FreqBoth – InfreqBoth × pronun × LangStatus	0.826	0.686	1.205	.228
	FreqNat – FreqBoth × pronun × LangStatus	0.048	0.589	0.081	.935
	InfreqNat – FreqNat × pronun × LangStatus	−0.192	0.736	−0.261	.794

Note: The two-level factors were coded with sum contrasts (i.e., pronunciation: non-target_pronun = −0.5/target_pronun = 0.5; LangStatus: Monolingual = −0.5/Bilingual = 0.5), while the factor phonotactic was coded with sliding contrasts (i.e., *InfreqBoth* vs. *FreqBoth*, *FreqBoth* vs. *FreqNative*, and *FreqNative* vs. *InfreqNative*). *: $p \leq .05$, **: $p \leq .01$, ***: $p \leq .001$.

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

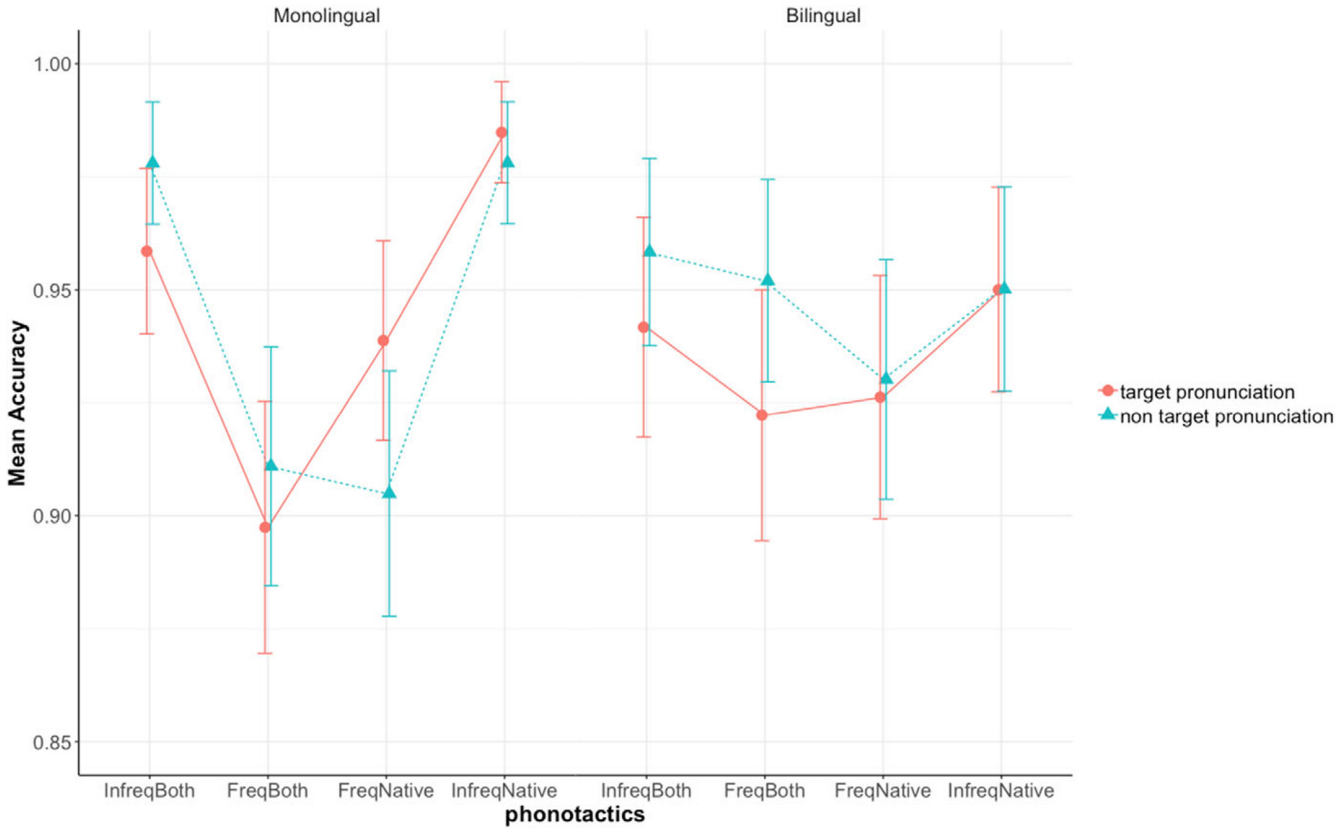


Figure 2. Summary of raw results from the lexical decision task, showing participants' mean accuracy (and CIs) in rejecting the nonwords only in relation to phonotactic condition, split based on language status (monolingual, bilingual) and language-specific pronunciation (non-target pronunciation, target pronunciation).

interaction between *FreqBoth-FreqNative* and pronunciation (p = .017); however, post-hoc comparisons failed to reveal significant accuracy differences between these two conditions in either pronunciation condition (stimuli pronounced by the speakers of the target [p = .791] or non-target language [p = .754]).

Because the task was twice as long for bilinguals (two blocks of 128 trials) compared to monolingually-raised participants (one block of 128 trials), we explored whether any phonotactic effect in the lexical decision task for the bilinguals could have emerged as a function of block. Hence, we ran an additional generalized mixed-effects model with the bilingual data only, with three independent variables: phonotactics and pronunciation in interaction, nested within block. There was a significant main effect of block (β = −0.953, SE = 0.190, z = −5.018, p < .001), indicating that participants' accuracy decreased between block1 and block2, and a significant main effect of *FreqBoth-InfreqBoth* in block2 (β = −0.789, SE = 0.377, z = −2.090, p = .037), indicating that accuracy was significantly higher in *InfreqBoth* (EMM = 0.958; 95%CI [0.907; 0.982]) compared to *FreqBoth* (EMM = 0.913; 95%CI [0.830; 0.959]; see Figure 3 and Table 4).

3.2.2. RT analyses

The results are presented in Table 5 and a plot summary of the raw data in Figure 4. They establish a significant main effect of *FreqBoth-InfreqBoth* (p = .041) and *InfreqNat-FreqNat* (p < .001) indicating that participants were faster in accurately deciding that a stimulus was not a word in their language in the *InfreqBoth* than *FreqBoth* conditions, and the *InfreqNat* than *FreqNat* conditions. While the main effect of language status was not significant (p = .153), it significantly interacted with phonotactics (*FreqBoth-InfreqBoth*: p = .002; *InfreqNat-FreqNat*: p = .001), indicating that phonotactic properties modulated participants' RTs differently between the monolingually-raised and bilingual groups. Post-hoc comparisons (the full output of orthogonal pairwise comparisons can be found in Supplementary Table S6) revealed that monolingually-raised participants' RTs significantly differed between *FreqBoth* (EMM = 514 ms, 95%CI [452 ms; 575 ms]) and *InfreqBoth* (EMM = 435 ms, 95%CI [371 ms; 498 ms]; z = 2.964, p = .009), and between *FreqNat* (EMM = 515 ms, 95%CI [453 ms; 577 ms]) and *InfreqNat* (EMM = 418 ms, 95%CI [359 ms; 477 ms]; z = −7.872, p < .001). There were no significant differences between these conditions in the bilingual participants (*FreqBoth*: EMM = 418 ms, 95%CI [330 ms; 507 ms], *InfreqBoth*: EMM = 389 ms, 95%CI [297 ms; 481 ms], z = 1.012, p = 0.630; *FreqNat*: EMM = 404 ms, 95%CI [311 ms; 496 ms], *InfreqNat*: EMM = 379 ms, 95%CI [291 ms; 467 ms], z = −1.534, p = 0.302). Finally, there was a significant interaction between pronunciation and *InfreqNat-FreqNat* (p = .002), indicating that, overall, the difference in RTs between these two phonotactic levels was bigger in the non-target pronunciation (*RT-diff* = −84.08 ms, SE = 12.6 ms, z = −6.653, p < .001) compared to the target pronunciation (*RT-diff* = −37.57 ms, SE = 12.6 ms, z = −2.989, p = .008). No other effects reached significance.

Just as for the accuracy analyses, we explored whether any phonotactic effect in the lexical decision task for the bilinguals could have emerged as a function of block. Hence, we ran an additional linear mixed-effects model with the bilingual data only, with three independent variables: phonotactics and pronunciation in interaction nested within block. There was a significant main effect of

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

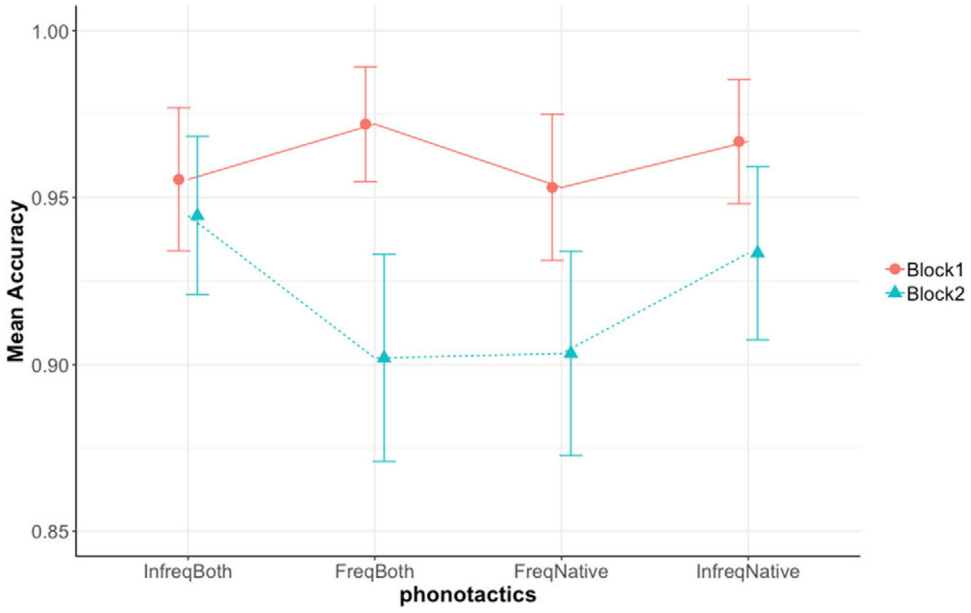


Figure 3. Summary of raw results from the lexical decision task, showing bilingual participants' mean accuracy (and CIs) for the nonwords only, in relation to phonotactic condition, split based on Block (Block1, Block2).

Table 4. Results of the exploratory generalized mixed-effects model assessing the effect of phonotactics and pronunciation within each block on simultaneous bilingual participants' accuracy in the lexical decision task (nonwords only)

	Parameter	Estimate	Std. Error	*z*	*p*
Effects	(Intercept)	3.921	0.355	11.046	<.001***
	Block	−0.953	0.190	−5.018	<.001***
	Block1: FreqBoth – InfreqBoth	0.551	0.482	1.142	.253
	Block2: FreqBoth – InfreqBoth	−0.789	0.377	−2.090	.037*
	Block1: FreqNat – FreqBoth	−0.565	0.470	−1.201	.230
	Block2: FreqNat – FreqBoth	−0.023	0.335	−0.069	.945
	Block1: InfreqNat – FreqNat	0.337	0.430	0.784	.433
	Block2: InfreqNat – FreqNat	0.608	0.325	1.868	.062
	Block1: pronunciation	0.076	0.302	0.252	.801
	Block2: pronunciation	0.408	0.224	1.824	.068
	Block1: FreqBoth – InfreqBoth × pronun	−0.093	0.884	−0.105	.916
	Block2: FreqBoth – InfreqBoth × pronun	0.520	0.648	0.802	.422
	Block1: FreqNat – FreqBoth × pronun	−0.997	0.871	−1.144	.252
	Block2: FreqNat – FreqBoth × pronun	−0.409	0.574	−0.713	.476
	Block1: InfreqNat – FreqNat × pronun	0.100	0.825	0.121	.904
	Block2: InfreqNat – FreqNat × pronun	−0.134	0.615	−0.218	.827

block ($\beta = -47.495$, SE = 8.286, $t = -5.732$, $p < .001$), indicating that participants' RTs decreased between block1 and block2, a significant main effect of *FreqNat-InfreqNat* in block2 ($\beta = -36.828$, SE = 16.650, $t = -2.212$, $p = .027$, Figure 5 and Supplementary Table S2), indicating that RTs were smaller in *InfreqNat* (EMM = 349 ms, 95%CI [277 ms; 421 ms]) compared to *FreqNat* (EMM = 386 ms, 95%CI [314 ms; 458 ms]), and a significant interaction between *FreqNat-InfreqNat* and pronunciation in block1, indicating that the phonotactic effect differed between pronunciation conditions in the first block. However, post-hoc analyses showed that the *InfreqNat-FreqNat* contrast was not significant in either the native ($p = .708$) or non-native ($p = .156$) condition when considered separately.

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

Table 5. Results of the mixed-effects model of corrected RTs in accurately rejecting nonwords in Experiment 2, taking into account the phonotactic conditions, language-specific pronunciation, and language status

	Parameter	Estimate	Std. Error	*t*	*p*
Effects	(Intercept)	433.895	26.605	16.309	<.001***
	FreqBoth – InfreqBoth	54.375	25.704	2.115	.041*
	FreqNat – FreqBoth	−6.455	21.721	−0.297	.768
	InfreqNat – FreqNat	−60.826	10.189	−5.970	<.001***
	pronunciation	−1.578	14.831	−0.106	.916
	LangStatus	−73.046	50.628	−1.443	.153
	FreqBoth – InfreqBoth × pronun	−1.131	41.961	−0.027	.979
	FreqNat – FreqBoth × pronun	30.253	37.152	0.814	.421
	InfreqNat – FreqNat × pronun	−46.517	14.836	−3.135	.002**
	FreqBoth – InfreqBoth × LangStatus	−49.449	22.305	−2.217	.030*
	FreqNat – FreqBoth × LangStatus	−15.978	16.661	−0.959	.339
	InfreqNat – FreqNat × LangStatus	71.807	20.368	3.525	<.001***
	pronun × LangStatus	−12.738	10.500	−1.213	.225
	FreqBoth – InfreqBoth × pronun × LangStatus	−19.412	29.767	−0.652	.514
	FreqNat – FreqBoth × pronun × LangStatus	38.890	30.011	1.296	.195
	InfreqNat – FreqNat × pronun × LangStatus	−13.335	29.644	−0.450	.653

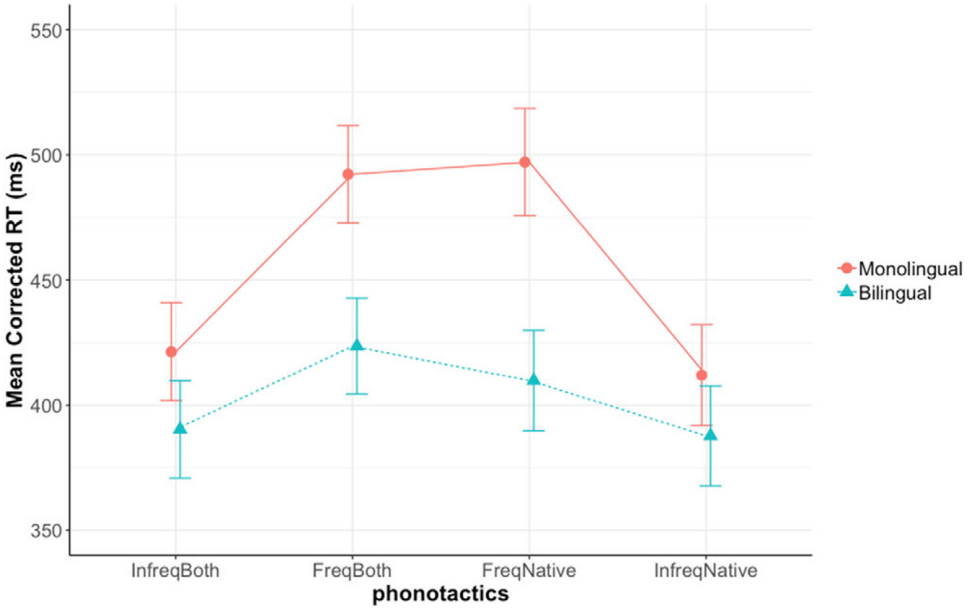


Figure 4. Summary of raw results from the lexical decision task, showing participants' mean corrected reaction times (and CIs) in accurately rejecting the nonwords in relation to phonotactic condition, split based on language status (monolingual, bilingual).

This pattern indicates that the interaction arises from a difference between conditions rather than from a robust effect within either condition alone.

3.3. Discussion

Experiment 2 probed simultaneous French–German bilinguals' phonotactic processing with a web-based lexical decision task, and compared their data to monolingually-raised participants. Overall, findings show that, as predicted, phonotactic properties modulated participants' lexical decisions: accuracy in rejecting a nonword as a real word of the native language was significantly better for nonwords composed of non-frequent compared to frequent regularities, entailing that the more frequent the phonotactic regularities, the more often a nonword was confused with a real French/German word. However, these effects were mainly driven

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

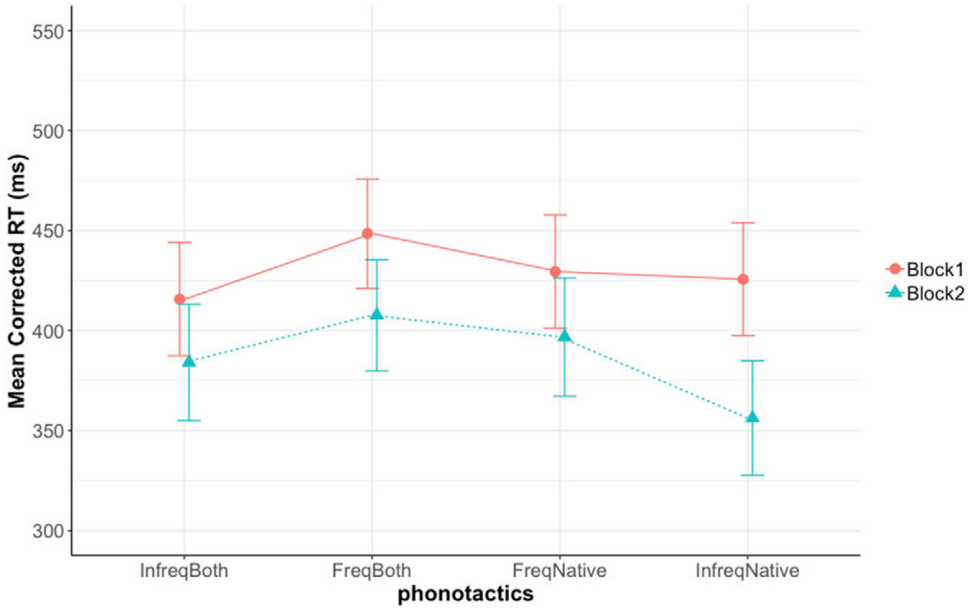


Figure 5. Summary of raw results from the lexical decision task, showing participants' mean corrected reaction times (and CIs) in accurately rejecting the nonwords in relation to phonotactic condition, split based on block (Block1, Block2).

by the monolingually-raised groups. Indeed, monolingually-raised and bilingual participants significantly differed in how phonotactic properties modulated their decisions: while significant effects of phonotactics on accuracy and RTs were found in monolingually-raised participants (in both the *InfreqBoth-FreqBoth* and *FreqNat-InfreqNat* levels), the same effects were not found in bilinguals. Only block analyses in bilinguals revealed phonotactic effects in the second half of the experiment, though restricted to the *FreqBoth-InfreqBoth* levels for accuracy, where the phonotactic properties of their two languages converged, and the *FreqNat-InfreqNat* levels for RTs, where they diverged.

These findings extend previous language-specific phonotactic processing effects in monolinguals (e.g., Vitevitch & Luce, 1999) to French and German, providing evidence that the task materials and web-based procedure were valid. In light of weak phonotactic effects in the bilingual group, the difference between monolingually-raised and bilingual participants' phonotactic processing suggests that bilinguals may rely differently on phonotactic properties during online processing than monolingually-raised individuals. One possible interpretation is that having to navigate between two phonotactic systems modulates how strongly language-specific phonotactic regularities are involved during lexical decisions. We return to these findings in the General Discussion.

4. General discussion

This study investigated phonotactic knowledge and processing in simultaneous French–German bilinguals and compared them to monolingually-raised speakers of these languages in two experiments: a metalinguistic well-formedness judgment task (Experiment 1) and a lexical decision task (Experiment 2). Because only one prior study assessed simultaneous bilinguals' phonotactics (Parlato-Oliveira et al., 2010), our first aim was to investigate whether, and if so how, their phonotactic knowledge and use of it during online processing would differ from two groups of monolingually-raised participants. Second, we investigated whether bilinguals' two phonotactic systems interact, potentially leading to differences from the monolingually-raised groups. To do so, we manipulated the phonotactic properties of our stimuli following Altenberg and Cairns (1983), such that nonwords started with word-initial clusters that were either frequent or non-frequent in both French and German (convergent phonotactics), or frequent in French but non-frequent in German or vice versa (divergent phonotactics), allowing us to directly assess whether phonotactic well-formedness in one language could influence bilinguals' behavior when processing the other. Furthermore, our stimuli were pronounced by both a native speaker of French and German, to explore the potential modulation of language interference by language-specific pronunciation factors.

Starting with the monolingually-raised individuals, our results in the two tasks demonstrated, as predicted, that phonotactic properties played a role in their performance. In the well-formedness judgment task, ratings were significantly higher for nonwords composed of frequent compared to non-frequent phonotactic regularities, such that more frequent phonotactic properties yielded higher wordlikeness ratings. In the lexical decision task, accuracy and RTs were significantly more accurate and faster for nonwords composed of non-frequent compared to frequent regularities. Results also showed that language-specific pronunciation modulated both monolingually-raised participants' overall ratings (with lower ratings in the non-target pronunciation condition) and the phonotactic effect in the judgment task, but not the lexical decision task, with a smaller but significant phonotactic effect when the items were pronounced by a speaker of the non-target language. This finding could imply that pronunciation properties have a strong influence on participants' wordlikeness judgments, and also that they interfere with the strength of phonotactic intuitions. However, methodological choices in Experiment 1 might have prompted such an effect of pronunciation. Indeed, it is possible that ratings were overall lower in the non-target pronunciation because of our task instructions: had we explicitly said to abstract away pronunciation cues, perhaps we would not have found such an effect. Furthermore, participants were presented first with

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

pronunciations of the target language followed by pronunciations of the non-target language, in a fixed order, which could have exacerbated differences in judgments between the two conditions. This might then also explain why pronunciation did not affect lexical decisions. Beyond the main effect of pronunciation, the fact that it influenced the magnitude of the phonotactic effect might result from more demanding processing for atypical pronunciations, with phoneme categorization being more impacted, which could have cascading effects impacting phonotactic and lexical processing (Floccia et al., 2009; Williams & Escudero, 2014). However, it should be noted that the effect of pronunciation was also found with the bilinguals. Thus, this finding should be interpreted with caution and be replicated in other experiments, as it did not emerge in the lexical decision task, possibly because processing at the lexical level is more resilient to sub-lexical pronunciation properties. However, task-specific effects could also explain such differences across the two experiments, since sub-lexical properties are not the explicit focus of the lexical decision task, and items pronounced by both speakers were presented together in a random order.

Turning to the simultaneous bilinguals, our main group of interest, we found a clear demonstration of their knowledge of the phonotactic properties of their two native languages in the well-formedness judgment task, extending to simultaneous bilingualism previous findings on sequential bilingualism (e.g., Altenberg & Cairns, 1983; Frisch & Brea-Spahn, 2010). However, findings from the lexical decision task suggest differences between bilingual and monolingually-raised individuals in their use of phonotactic knowledge during processing.

Their performance in the well-formedness judgment task suggests that indeed, simultaneous bilinguals apply phonotactic knowledge appropriately to the language being processed even when asked to shift between their languages within the same task. This suggests that bilinguals were operating in language-specific modes in this rating task, and that they possess detailed phonotactic knowledge in both of their languages. This is further supported by the fact that pronunciation interfered with phonotactic wordlikeness in the same way in the bilingual and monolingually-raised groups, even though the former group should have more experience with both pronunciations. Nevertheless, ratings in bilinguals were not modulated by phonotactic properties in the exact same way as in monolingually-raised participants, as revealed by one (and only one) difference between groups, namely a significantly smaller difference in ratings between the *FreqBoth-InfreqBoth* levels in the bilingual compared to the monolingually-raised group. However, this difference did not extend to the *FreqNat-InfreqNat* levels, suggesting that this effect is small and not robust. Moreover, the fact that such a difference emerged only in levels for which the phonotactic properties converged across French and German (i.e., regularities were either frequent or non-frequent in both languages) and not in levels where the two languages diverged phonotactically makes it unlikely to result from language-nonselective activation of bilinguals' phonotactic representations. Rather, it is possible that it results from a group difference in acceptability thresholds, with simultaneous bilinguals being more willing to accept any sequence of sounds as new words in their languages, perhaps because they are more regularly exposed to language variations at various levels (see Bialystok, 2001; for a review of metalinguistic flexibility in bilingual children).

Findings in the lexical decision task, which evaluates phonotactic processing rather than knowledge, suggest a completely different picture, showing clear differences between the bilingual and monolingually-raised participants. Whereas monolingually-raised participants' accuracy and RTs significantly differed as a function of phonotactics (in both the *InfreqBoth-FreqBoth* and *FreqNat-InfreqNat* levels), no significant overall phonotactic effects were found in the bilinguals. However, an additional analysis comparing bilinguals' performance between the two blocks of the experiment (each block corresponding to one of their languages) showed that bilinguals' performance was modulated by phonotactics only in the second block, with a significant difference in accuracy between the levels *FreqBoth-InfreqBoth* and a significant difference in RTs between the levels *FreqNat-InfreqNat.* These effects might reflect both task-specific factors and interactions between bilinguals' two phonotactic systems. In the first block, the lack of a significant phonotactic effect may relate to ceiling-level accuracy and relatively slow decisions, while in the second block, participants' accuracy and RTs significantly decreased (e.g., for accuracy, from 96% to 92%; for RTs, from 430 to 386 ms). Because stimuli in both pronunciations occurred within a block, it might have been harder to efficiently process the stimuli's phonotactics compared to Experiment 1, requiring bilinguals to make more effort to suppress the activation of words in the other language and prompting them to be particularly careful in their decisions in the first half of the experiment to cope with such unexpected variation. It is only once familiarized with the task that some effects emerged, which coincided with more errors and quicker responses. Nevertheless, the significant phonotactic effects found in the second block differed between accuracy and RT analyses. For accuracy, our results would fit with a language nonselective activation account, which might have modulated the phonotactic effect: the smaller, non-significant, difference between the *FreqNat* and *InfreqNat* levels could result from interference in phonotactic processing when the phonotactic properties diverged in the bilinguals' two native languages, while the bigger, significant, difference between the *FreqBoth* and *InfreqBoth* levels could result from a strengthened phonotactic effect when the phonotactic properties converged in their two native languages. If our interpretation is correct, this effect would (weakly) replicate, and extend to simultaneous bilinguals, what Altenberg and Cairns (1983) found with sequential German–English bilinguals using written stimuli: in their study, sequential bilinguals were slower than English monolinguals to reject nonwords as real words of English when these nonwords were composed of legal phonotactic regularities in German that were illegal in English, suggesting interference of German phonotactics when processing English. The fact that in our study such effects emerged only in lexical decisions but not well-formedness judgments is consistent with the idea that online processing tasks are more sensitive to cross-language interactions than metalinguistic tasks (Freeman et al., 2022). However, this effect only emerged in the second half of Experiment 2 and was not confirmed in the RT analyses. In the RT analysis, the only significant effect emerging in block2 was specific to the *FreqNat-InfreqNat* levels, with quicker decisions in *InfreqNat* compared to *FreqNat.* Thus, the language nonselective interpretation of bilinguals' phonotactic processing should be taken cautiously and replicated before strong conclusions are drawn about the mechanisms underlying such group differences. Alternatively, it is also possible that co-activation of their two phonotactic systems is reduced in simultaneous bilinguals, as opposed to sequential bilinguals, because of smaller asymmetries in proficiency/knowledge between their two languages (Sebastián-Gallés et al., 2005; but see Luk & Bialystok, 2013), or because of smaller L1 interference effects (Flege, 2003; Iverson et al., 2003). If so, L1 interference during L2 processing might be weaker in

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

bilingually-raised individuals. Studies directly comparing these populations will be important to clarify this issue.

In summary, Experiment 1 suggests that simultaneous bilinguals have knowledge of the phonotactics of both native languages comparable to that of monolingually-raised individuals and can use such knowledge offline in a language-specific way. Moreover, Experiment 2 suggests that bilinguals process phonotactics differently from monolingually-raised individuals, possibly due to cross-language interactions during online processing, although this requires further investigation.

Before concluding, we would like to describe some limitations of the current study. One concerns the language background of the monolingually-raised control groups. Participants were required to have had exposure to only one language during early childhood; however, we did not collect information about languages learned later in life. Because foreign-language instruction is widespread from primary school onward in the French and German educational systems, it is likely that many participants had some classroom-based knowledge of foreign language (most commonly English or Spanish). Importantly, such experience differs substantially from early simultaneous bilingual exposure and does not entail relatively balanced use of two phonotactic systems from infancy. Furthermore, if any participants had acquired knowledge of the other target language, this would likely have reduced differences between monolingually-raised and bilingual groups, making our comparison more conservative. Therefore, we acknowledge that systematically documenting later-acquired language experience in our control groups would have further informed group comparisons. A second limitation relates to the cross-task comparison design implemented to investigate bilinguals' phonotactic knowledge (Experiment 1) and processing (Experiment 2) and how the two relate. The two tasks differed in their structure and dependent measures: the judgment task used a Likert scale, whereas the lexical decision task relied on accuracy/RTs. These differences in response measures might influence statistical sensitivity and thus limit the strength of conclusions stemming from direct comparisons across the two tasks. Nevertheless, we believe that by clearly defining the construct investigated with each task (knowledge vs. processing), we refrained from making direct comparisons. A third limitation concerns the pronunciation manipulation, as only one speaker per language was used to pronounce our items, raising the possibility that some effects result from idiosyncratic speaker characteristics. However, the pattern of results was symmetric across language groups suggesting that the findings are unlikely to be driven only by speaker-specific properties (e.g., in Experiment 1, French monolingually-raised participants showed reduced effects for German pronunciations and vice versa, although this is not explicitly apparent in our models because the two monolingually-raised groups were collapsed together). Finally, a fourth limitation relates to the web-based nature of the present study. A clear advantage of such a design is the facilitation of recruitment and testing of uncommon populations (in our case, simultaneous French–German bilinguals). A clear disadvantage is that the lack of experimental control over both participants (e.g., mental states, settings, and levels of external distraction during the experiment) and technical equipment often results in noisier data, which might have reduced statistical power to reliably detect some effects. While our power analyses indicated that sample sizes were sufficiently large to detect phonotactic effects in both tasks, the data might have been underpowered for RT analyses in Experiment 2, which is why only accuracy analyses were preregistered. Replication under controlled laboratory settings would therefore be useful.

In conclusion, our study showed, with two tasks, that individuals raised in a bilingual French–German environment (i.e., simultaneous bilinguals) have *knowledge* of the phonotactic properties of their two languages at similar level as monolingually-raised individuals, and are able to activate such knowledge selectively. Nevertheless, our findings suggest that bilinguals' *phonotactic processing* differs from that of monolingually-raised individuals in their two native languages, possibly because of co-activation of phonotactic representations in their two languages, resulting in interference when properties conflict across languages. Our study opens up the question of whether simultaneous bilinguals differ from sequential bilinguals in the degree of language-selective activation of phonotactics in their two languages. While most studies showed an important effect of L1 phonotactic knowledge during L2 processing in sequential bilinguals (but see Altenberg & Cairns, 1983 for the opposite direction), cross-language effects were less clear and robust in our study, which could imply greater abilities in navigating between their two languages' phonotactics in simultaneous bilinguals.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728926101667.

Data availability statement. The data and the R-script for data analysis are available on the following open repository in OSF: https://osf.io/kqu48/overview?view_only=8786b2b306e44ccfb066d96eb2a20943

Acknowledgements. We thank all the persons who participated in the study. A special thanks to Ekaterina Artemewa, Vladislav Trakal, and Yulia Pachenko for their precious help in conducting this study.

Author contribution. Conceptualization: L.P., T.N., N.B.-A.; Data curation: L.P.; Formal analysis: L.P., T.N., N.B.-A.; Funding acquisition: T.N., N.B.-A; Investigation: L.P.; Methodology: L.P., T.N., N.B.-A.; Project administration: L.P., T.N., N.B.-A.; Resources: T.N., N.B.-A.; Supervision: T. N., N.B.-A; Validation: T.N., N.B.-A.; Visualization: L.P.; Writing–original draft & revision: L.P.; Writing– revision & editing: T.N., N.B.-A.

Funding statement. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317633480 – SFB 1287, Project C07; and by a public grant overseen by the IdEx Université Paris Cité (ANR-18-IDEX-0001) as part of the Labex Empirical Foundations of Linguistics – EFL.

Competing interests. The authors declare none.

Ethical statement. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Altenberg, E. P., & **Cairns, H. S.** (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, **22**(2), 174–188. https://doi.org/10.1016/S0022-5371(83)90134-2

Anisfeld, M., **Anisfeld, E.**, & **Semogas, R.** (1969). Cross-influences between the phonological systems of Lithuanian-English bilinguals. *Journal of Verbal Learning and Verbal Behavior*, **8**(2), 257–261. https://doi.org/10.1016/S0022-5371(69)80072-1

Baayen, R. H., & **Milin, P.** (2010). Analyzing reaction times. *International Journal of Psychological Research*, **3**(2), 12–28. https://doi.org/10.21500/20112084.807

Bailey, T. M., & **Hahn, U.** (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, **44**(4), 568–591. http://doi.org/10.1006/jmla.2000.2756

Barr, D. J., **Levy, R.**, **Scheepers, C.**, & **Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., **Mächler, M.**, **Bolker, B.**, & **Walker, S.** (2015). Fitting linear mixed-effects models Usinglme4. *Journal of Statistical Software*, **67**(1). https://doi.org/10.18637/jss.v067.i01

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, **21**, 169–181. https://doi.org/10.1017/S0267190501000101

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, **5**(9), 341–345.

Boll-Avetisyan, N. (2011). Second language probabilistic phonotactics and structural knowledge in short-term memory recognition. In N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th annual Boston University conference on language development* (Vol. **1**, pp. 60–72). Cascadilla Press.

Boll-Avetisyan, N., **Bhatara, A.**, **Unger, A.**, **Nazzi, T.**, & **Höhle, B.** (2020). Rhythmic grouping biases in simultaneous bilinguals. *Bilingualism: Language and Cognition*, **23**(5), 1070–1081. https://doi.org/10.1017/S1366728920000140

Boll-Avetisyan, N., & **Kager, R.** (2016). Is speech processing influenced by abstract or detailed phonotactic representations? The case of the obligatory contour principle. *Lingua*, **171**, 74–91. https://doi.org/10.1016/j.lingua.2015.11.008

Brysbaert, M. (2013). LexTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, **53**(1), 23–37. https://doi.org/10.5334/pb-53-1-23

Carlson, M. T., **Goldrick, M.**, **Blasingame, M.**, & **Fink, A.** (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition*, **19**(5), 939–954. https://doi.org/10.1017/S1366728915000334

Christensen, R. H. B. (2019). *Ordinal—regression models for ordinal data. R package version 2019.12–10.* https://CRAN.R-project.org/package=ordinal

De Groot, A. M., & **Nas, G. L.** (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, **30**(1), 90–123. https://doi.org/10.1016/0749-596X(91)90012-9

Dumay, N., **Frauenfelder, U. H.**, & **Content, A.** (2002). The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language*, **81**(1–3), 144–161. https://doi.org/10.1006/brln.2001.2513

Dupoux, E., **Parlato, E.**, **Frota, S.**, **Hirose, Y.**, & **Peperkamp, S.** (2011). Where do illusory vowels come from? *Journal of Memory and Language*, **64**(3), 199–210. https://doi.org/10.1016/j.jml.2010.12.004

Dupoux, E., **Peperkamp, S.**, & **Sebastián-Gallés, N.** (2010). Limits on bilingualism revisited: Stress 'deafness' in simultaneous French–Spanish bilinguals. *Cognition*, **114**(2), 266–275. https://doi.org/10.1016/j.cognition.2009.10.001

Finger, H., **Goeke, C.**, **Diekamp, D.**, **Standvoß, K.**, & **König, P.** (2017). LabVanced: A unified JavaScript framework for online studies. In *International Conference on Computational Social Science*. Cologne. https://www.labvanced.com/content/research/en/publications/

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. Schiller & A. Meyer (Eds.), *Phonetics and phonology in language comprehension and production. Differences and similarities* (Vol. **6**, pp. 319–355). De Gruyter Mouton.

Floccia, C., **Butler, J.**, **Goslin, J.**, & **Ellis, L.** (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, **38**(4), 379–412. https://doi.org/10.1007/s10936-008-9097-8

Freeman, M. R., **Blumenfeld, H. K.**, **Carlson, M. T.**, & **Marian, V.** (2022). First-language influence on second language speech perception depends on task demands. *Language and Speech*, **65**(1), 28–51. https://doi.org/10.1177/0023830920983368

Freeman, M. R., **Blumenfeld, H. K.**, & **Marian, V.** (2016). Phonotactic constraints are activated across languages in bilinguals. *Frontiers in Psychology*, **7**, 702. https://doi.org/10.3389/fpsyg.2016.00702

Frisch, S. A., & **Brea-Spahn, M. R.** (2010). Metalinguistic judgments of phonotactics by monolinguals and bilinguals. *Laboratory Phonology*, **1**(2), 345–360. http://doi.org/10.1515/labphon.2010.018

Frisch, S. A., **Large, N. R.**, **Zawaydeh, B.**, & **Pisoni, D. B.** (2001). Emergent phonotactic generalizations in english and arabic. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. John Benjamins.

Hallé, P. A., **Dominguez, A.**, **Cuetos, F.**, & **Segui, J.** (2008). Phonological mediation in visual masked priming: Evidence from phonotactic repair. *Journal of Experimental Psychology*, **34**(1), 177–192. https://doi.org/10.1037/0096-1523.34.1.177

Hallé, P. A., **Segui, J.**, **Frauenfelder, U.**, & **Meunier, C.** (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, **24**(2), 592–608. https://doi.org/10.1037/0096-1523.24.2.592

Hallé, P. A., & **Best, C. T.** (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/clusters. *The Journal of the Acoustical Society of America*, **121**(5), 2899–2914. https://doi.org/10.1121/1.2534656

Hay, J., **Pierrehumbert, J.**, & **Beckman, M.** (2004). Speech perception, well-formedness and the statistics of the lexicon. In J. Local, R. Ogden, R. Temple, J. Local, R. Ogden, & R. Temple (Eds.), *Papers in laboratory phonology vi*. Cambridge University Press.

Henninger, F., **Shevchenko, Y.**, **Mertens, U. K.**, **Kieslich, P. J.**, & **Hilbig, B. E.** (2020). lab.js: A free, open, online study builder. https://doi.org/10.5281/zenodo.597045

Iverson, P., **Kuhl, P. K.**, **Akahane-Yamadac, R.**, **Dieschd, E.**, **Tohkurae, Y.**, **Kettermannf, A.**, & **Siebert, C.** (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, **87**(1), B47–B57. https://doi.org/10.1016/S0010-0277(02)00198-1

Kuznetsova, A., **Brockhoff, P. B.**, & **Christensen, R. H.** (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lagrou, E., **Hartsuiker, R. J.**, & **Duyck, W.** (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **37**(4), 952–965. https://doi.org/10.1037/a0023217

Lagrou, E., **Hartsuiker, R. J.**, & **Duyck, W.** (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition*, **16**(3), 508–517. https://doi.org/10.1017/S1366728912000508

Lemhöfer, K., & **Broersma, M.** (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, **44**(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lenth, R. (2024). emmeans: Estimated marginal means, aka least-squares means. R package version 1.10.6-090002. https://rvlenth.github.io/emmeans/

Lentz, T. O., & **Kager, R. W.** (2015). Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech*, **58**(3), 387–413. https://doi.org/10.1177/0023830914559572

Liddell, T. M., & **Kruschke, J. K.** (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, **79**, 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

Luk, G., & **Bialystok, E.** (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, **25**(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

Marian, V., **Blumenfeld, H. K.**, & **Kaushanskaya, M.** (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, **50**(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067

Marian, V., & **Spivey, M.** (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, **6**(2), 97–115. https://doi.org/10.1017/S1366728903001068

Massaro, D. W., & **Cohen, M. M.** (1983). Phonological context in speech perception. *Perception & Psychophysics*, **34**(4), 338–348. https://doi.org/10.3758/BF03203046

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, **39**(1), 21–46. https://doi.org/10.1006/jmla.1998.2568

Mersad, K., & **Nazzi, T.** (2011). Transitional probabilities and positional frequency phonotactics in a hierarchical model of speech segmentation. *Memory & Cognition*, **39**(6), 1085–1093. https://doi.org/10.3758/s13421-011-0074-3

Oh, Y., **Todd, S.**, **Beckner, C.**, **Hay, J.**, **King, J.**, & **Needle, J.** (2020). Non-Māori-speaking New Zealanders have a Māori proto-lexicon. *Scientific Reports*, **10**(1), 22318–22319. http://doi.org/10.1038/s41598-020-78810-4

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press

Otake, T., **Yoneyama, K.**, **Cutler, A.**, & **Van Der Lugt, A.** (1996). The representation of Japanese moraic nasals. *The Journal of the Acoustical Society of America*, **100**(6), 3831–3842. https://doi.org/10.1121/1.417239

Parlato-Oliveira, E., **Christophe, A.**, **Hirose, Y.**, & **Dupoux, E.** (2010). Plasticity of illusory vowel perception in Brazilian-Japanese bilinguals. *The Journal of the Acoustical Society of America*, **127**(6), 3738–3748. https://doi.org/10.1121/1.3327792

Pitt, M. A., & **McQueen, J. M.** (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, **39**(3), 347–370. https://doi.org/10.1006/jmla.1998.2571

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richtsmeier, P. T. (2011). Word-types not word-tokens, facilitate extraction of phonotactic sequences by adults. *Laboratory Phonology*, **2**(1), 157–183. http://doi.org/10.1515/labphon.2011.005

Schulpen, B., **Dijkstra, T.**, **Schriefers, H. J.**, & **Hasper, M.** (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **29**(6), 1155–1178. https://doi.org/10.1037/0096-1523.29.6.1155

Sebastián-Gallés, N., & **Bosch, L.** (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance*, **28**(4), 974–989. http://doi.org/10.1037/0096-1523.28.4.974

Sebastián-Gallés, N., **Echeverría, S.**, & **Bosch, L.** (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, **52**(2), 240–255. https://doi.org/10.1016/j.jml.2004.11.001

Smith, K. L., & **Pitt, M. A.** (1999). Phonological and morphological influences in the syllabification of spoken words. *Journal of Memory and Language*, **41**(2), 199–222. https://doi.org/10.1006/jmla.1999.2646

Trapman, M., & **Kager, R.** (2009). The Acquisition of Subset and Superset Phonotactic Knowledge in a second language. *Language Acquisition*, **16**(3), 178–221. https://doi.org/10.1080/10489220903011636

Van Assche, E., **Brysbaert, M.**, & **Duyck, W.** (2020). Bilingual lexical access. In R. R. Heredia, & A. B. Cieślicka (Eds.),*Bilingual lexical ambiguity resolution* (pp. 42–67). Cambridge University Press.

Van der Lugt, A. H. (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics*, **63**(5), 811–823. https://doi.org/10.3758/BF03194440

Vitevitch, M. S., & **Luce, P. A.** (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, **9**(4), 325–329. http://doi.org/10.1111/1467-9280.00064

Vitevitch, M. S., & **Luce, P. A.** (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, **40**(3), 374–408. http://doi.org/10.1006/jmla.1998.2618

Vitevitch, M. S., & **Luce, P. A.** (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, **52**(2), 193–204. https://doi.org/10.1016/j.jml.2004.10.003

Vitevitch, M. S., **Luce, P. A.**, **Charles-Luce, J.**, & **Kemmerer, D.** (1996). Phonotactic and metrical influences on adult ratings of spoken nonsense words. In *Proceeding of fourth international conference on spoken language processing* (Vol. **1**, pp. 82–85). IEEE. https://doi.org/10.1109/ICSLP.1996.607036

Weber, A., & **Cutler, A.** (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America*, **119**(1), 597–607. https://doi.org/10.1121/1.2141003

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer International Publishing. https://ggplot2.tidyverse.org

Williams, D., & **Escudero, P.** (2014). Native and non-native speech perception. *Acoustics Australia*, **42**(2), 379–383.

https://doi.org/10.1017/S1366728926101667 Published online by Cambridge University Press