



OPEN

Music, Language, and The N400: ERP Interference Patterns Across Cognitive Domains

Nicole Calma-Roddin^{1,3}✉ & John E. Drury²

Studies of the relationship of language and music have suggested these two systems may share processing resources involved in the computation/maintenance of abstract hierarchical structure (syntax). One type of evidence comes from ERP interference studies involving concurrent language/music processing showing interaction effects when both processing streams are simultaneously perturbed by violations (e.g., syntactically incorrect words paired with incongruent completion of a chord progression). Here, we employ this interference methodology to target the mechanisms supporting long term memory (LTM) access/retrieval in language and music. We used melody stimuli from previous work showing out-of-key or unexpected notes may elicit a musical analogue of language N400 effects, but only for familiar melodies, and not for unfamiliar ones. Target notes in these melodies were time-locked to visually presented target words in sentence contexts manipulating lexical/conceptual semantic congruity. Our study succeeded in eliciting expected N400 responses from each cognitive domain independently. Among several new findings we argue to be of interest, these data demonstrate that: (i) language N400 effects are delayed in onset by concurrent music processing only when melodies are familiar, and (ii) double violations with familiar melodies (but not with unfamiliar ones) yield a sub-additive N400 response. In addition: (iii) early negativities (RAN effects), which previous work has connected to musical syntax, along with the music N400, were together delayed in onset for familiar melodies relative to the timing of these effects reported in the previous music-only study using these same stimuli, and (iv) double violation cases involving unfamiliar/novel melodies *also* delayed the RAN effect onset. These patterns constitute the first demonstration of N400 interference effects across these domains and together contribute previously undocumented types of interactions to the available pool of findings relevant to understanding whether language and music may rely on shared underlying mechanisms.

Over the past decades a fruitful body of research examining language and music processing has shown we have at least as much to learn from joint investigations of their relationship as we do from studying either in isolation¹⁻⁵. One such line of investigation concerns the extent to which the processing of language and music may be subserved by shared brain mechanisms or whether these systems are rather supported by independent modular systems^{4,6-10}. Consider in this connection the notion that linguistic and musical cognition may be supported by shared syntactic processing resources^{1,8,9,11,12}. Accounts of these domains as cognitive systems each require reference to sequentially and hierarchically organized internal representations (i.e., a *syntax*¹³⁻¹⁶). Proponents of shared syntactic processing maintain that while language/music involve different constituent vocabulary (i.e., words, morphemes vs. notes, chords, etc.), complex structured representations in these domains may be supported by some shared or overlapping set of processing resources.

In the present study we pose the question of whether other components of language and music processing may be shared, in particular the mechanisms responsible for managing access/retrieval of information from long term memory (LTM). To this end, we pursued an approach drawing upon previous work probing language and music syntactic processing with interference paradigms¹⁷⁻¹⁹, in particular experiments that measure event related potentials (ERPs^{12,20}). Such studies typically involve simultaneous processing of language and music stimuli, for example, word-by-word sentence reading while listening to chord progressions (or melodies) with the presentation of

¹Department of Behavioral Sciences, New York Institute of Technology, Old Westbury, New York, USA. ²School of Linguistic Sciences and Arts, Jiangsu Normal University, Xuzhou, China. ³Department of Psychology, Stony Brook University, New York, USA. ✉e-mail: ncalmaro@nyit.edu

each chord (or note) time-locked to the presentation of each word. Coupled with electrophysiological measurements, such studies make it possible to observe brain response profiles to manipulations disrupting either the music or the language processing streams independently, or both together (double violations). Interactions that arise in such paradigms have been interpreted as indexing simultaneous demands imposed by the two processing streams upon some common set of underlying neural resources.

Importantly, ERP studies that have investigated syntactic violations within each domain *separately* have demonstrated response profiles that are at least superficially similar. In language, disruption of syntactic processing with violations involving major word category (e.g., encountering a verb in a context where only nouns are possible) or incorrect inflectional/morphological marking (see reviews^{21,22}), often elicit a biphasic response consisting of an anterior negative-going deflection (“left anterior negativities” or “LANs”) followed by a relative positivity over posterior regions (P600s). In music, similar biphasic violation responses involving anterior negativities and P600s can be elicited by disruptions of harmonic processing (e.g., out-of-key notes in melodies or incongruent continuations of chord progressions^{12,20,23–25}). The musical correspondent of linguistic LAN effects are referred to as “RANs” (for “right anterior negativities”), given that the first report of these effects demonstrated a right lateralized scalp distribution²³ (note we do not, in the present report, distinguish between earlier (“ERAN”) versus later onset RAN-type effects; our view is that there is no such functional distinction, similar to the analogous effects in the language domain; see Steinhauer & Drury²²).

Although the mere similarity of linguistic LAN and musical RAN does not make a case for overlap of underlying neural systems, when music and language syntactic violations are encountered together in interference paradigms, linguistic LAN effects have been shown to be attenuated¹². Furthermore, available data suggest this type of ERP interaction profile is specific to simultaneous *syntactic* disruptions. For example, in Koelsch *et al.*¹² ERP differences were also compared between low and high probability sentence-final words, a manipulation introduced to elicit a language N400 response linked to expectancy/congruency of incoming words (see Kutas & Federmeier²⁶ and below). Interestingly, the language N400 response appears *not* to be influenced by concurrent music syntactic violations^{12,20}, (a finding that we replicate here with unfamiliar melodies – see below).

The target of the present investigation is the nature of the relationship between language and music long term memory access/retrieval mechanisms^{27–29} as viewed through the lens of the N400 ERP response^{26,30}. The motivation for our study relates to the apparent selective influence of music syntactic violations on linguistic LAN effects, while language N400 responses seem immune to such disturbance^{12,20}. Here we pose the question of whether this language/music dissociation would hold in circumstances where N400 effects are elicited by both processing streams simultaneously. Note that while we presuppose a particular view of the N400 as indexing semantic LTM access/retrieval mechanisms, the work reported here is ultimately neutral on this topic and could be reinterpreted under other conceptions of this ERP response (we reiterate this point below, and see our Discussion).

N400 effects in music processing? In contrast to linguistic LAN effects, the N400 is a negative-going ERP component with a broad scalp distribution, typically maximal in amplitude over central/parietal electrodes, which peaks around 400 ms post-stimulus onset. In language ERP research this component has been argued to reflect the access/retrieval (or binding) of conceptual semantic information stored in long term memory²⁶. Alternatively, the N400 has been discussed as indexing semantic integration^{31,32} or inhibition³³ (and, as just noted above, should either of these alternative conceptions turn out to be correct, this would change how we interpret the results reported here, but this does not affect the core empirical findings; see Discussion).

Though the initial discovery of the N400 was linked to visual language processing³⁰, subsequent research has revealed a wider family of N400-like effects that can be elicited across modalities and can be triggered in connection with a fairly breath-taking variety of different types of stimulus manipulations across cognitive domains^{26,34}. This makes the N400 an ideal target for studying a wide array of issues in cognitive processing, but this also exposes a curious gap when we consider music cognition. Though the N400 effect has been leveraged in previous studies to examine associative priming between language and music^{33–38}, to our knowledge there is only *one* report documenting an N400-like effect elicited by manipulations of auditorily presented music stimuli *in isolation*²⁴ (though see other possible cases^{39,40}).

Moreover, even in the lone study reporting a music related N400 response²⁴, the interpretation of the observed ERP pattern as a *bona fide* N400 is significantly hedged (for reasons we will canvas below). Before turning to that study, note that while previous work makes it clear that N400s *can* be elicited by music stimuli, it is important that such findings have involved cross-domain associative and affective priming^{33–38}. While such findings are significant, as they make it clear that music stimuli can be linked to (arguably abstract, amodal) conceptual representations, it is not obvious how this could be employed to create a *violation paradigm* to examine cross-domain interference between language and music involving the neural mechanisms underlying the N400 (similar to the work on syntax mentioned above).

This leads us to Miranda and Ullman (2007)²⁴, (henceforth: “M&U”), where ERP responses were examined for out-of-key notes in both familiar and matched unfamiliar melodies, and for incorrect but in-key notes in familiar melodies. As the present study makes use of the same music stimuli employed by M&U (see §4), we summarize their findings in detail as we expected (most of) their reported patterns to replicate.

First, M&U found that out-of-key notes in both melody types elicited RANs, P600s, and late anterior negativities (interpreted as so-called N500 effects, claimed to index processing of “intramusical meaning”^{41,42}, but see Featherstone *et al.*⁴³). Second, only familiar melodies elicited a frontal positivity, interpreted as a P3a effect (indexing attention capture⁴⁴), and the familiar melody P600 violation effect was larger than it was for unfamiliar melodies (replicating earlier findings³⁹). Third, and most importantly, a posterior negativity for familiar melody violations *only* was also reported, which (for the out-of-key condition) overlapped in timing with the RAN effect but was slightly later in onset. This posterior negativity also manifested for incorrect but in-key notes in the

familiar melodies. The scalp distribution of the posterior negativity, coupled with the fact it was elicited only for unexpected notes (whether in-key or out-of-key) in melodies that must be encoded in long term memory, drove M&U to label this negativity, tentatively, as an N400 (which they append with a ?-mark).

Interpreting their “N400?” as a *bona fide* species of N400 makes sense: in the case of familiar melodies, the processes involved in access/retrieval of long term memory representations could be disturbed by unpredicted incongruent continuations in ways that may be analogous to what happens in language when semantically unexpected or incongruent words are encountered. This could be taken to be consistent with some findings in neuro-imaging studies that have jointly investigated semantic memory in language and music, where some have argued for partial overlap of underlying networks across these and other cognitive domains^{27–29}.

However, inspection of M&U’s data, in combination with previous findings that failed to elicit music N400 responses^{45,46}, leads one to be sympathetic with the hesitancy exhibited in M&U’s interpretation of this response pattern. While their “N400?” is consistent across their in-key and out-of-key familiar melody violations, the effect did not exhibit the classic centro-parietal scalp distribution typical of the N400, but rather manifested a fairly restricted parietal/occipital distribution. This may have been due to the fact that this effect coincided with the large amplitude anterior P3a, perhaps canceling out what would otherwise have been a more prototypical N400 distribution. However, M&U’s follow-up analyses examining musicians/non-musicians separately showed that the musicians were driving the P3a effect, which was attenuated in the non-musicians. If the P3a was cancelling a music N400 effect that would otherwise have exhibited a broader distribution, we might have expected it to show just such a broader pattern in the non-musicians (where the P3a was absent). Instead, both cases of the “N400?” exhibited the more restricted parietal/occipital distribution, suggesting two possibilities: first, less musical expertise may result in *both* the absence of the P3a *and* a smaller, more restricted music N400 effect. A second possibility is that M&U’s “N400?” is simply differently distributed than linguistic N400 effects. After all, N400-like responses in other nonlinguistic domains^{26,34,47} can exhibit scalp distributions different from the canonical cases of the language N400 as well.

In any case, M&U’s study is *unique* in having documented a possible music analogue of the linguistic N400 that may be independent of associative or affective cross-domain priming. However, it is important to note that many of the melodies used by M&U had lyrics associated with them. This raises the question of whether melody familiarity is, on its own, sufficient to elicit their N400? effect, or whether this finding may be restricted to cases where melodies are stored in long term memory together with linguistic information (lyrics). However, whether this is the case or not is an issue that is secondary to the main goals of the present study. The aim of our work is to determine whether the posterior negativity seen for familiar melodies is indeed best understood as belonging in the N400 family, and whether simultaneous elicitation of such responses across language and music will demonstrate interference profiles similar to what has been previously shown for syntax. Showing this to be the case would be significant, opening up a range of possible follow-up work examining further details of such response patterns in ways that could shed some light both on the nature of the N400 and the underlying mechanisms that it is thought to index. Further, documenting interference effects for this component would contribute an important and previously undocumented type of interaction to the inventory of findings relevant to understanding how language and music may rely on shared underlying neurocognitive mechanisms.

Present study. We reemployed M&U’s familiar/unfamiliar melody stimuli in a language/music interference paradigm. Our goal was to both separately and simultaneously elicit music and language N400 effects in a within-participants experimental design, and examine the question of whether these responses might show interaction profiles analogous to those previously reported for manipulations that trigger linguistic LAN and music RAN effects. More generally, for many (if not most) instances where N400-like effects have been reported across different cognitive domains, it has yet to be shown that the underlying systems involved are the same, as opposed to functionally similar but nonetheless distinct neural systems⁴⁷.

To accomplish this, following previous interference studies of language/music syntax, we constructed sentences manipulating conceptual semantic congruency at the VERB position in relative clauses (e.g., ...*the ball that the boy has KICKED/*BAKED*). These contexts for our critical target words were chosen so that we could test these same materials in a second experiment inducing morpho-syntactic violations (e.g., ...*the ball that the boy will *KICKED*...). (Note: The present study constitutes the first stage of a larger scale multi-part investigation. The syntactic violations follow-up experiment, and between-group comparisons to single modality effects elicited by the language and music stimuli employed here, as well as an investigation of individual differences and item analyses probing familiarity, will be detailed in a longer forthcoming report.)

These target noun-phrase + relative clause fragments were aligned to musical notes in familiar and unfamiliar melodies, using (a subset of) the stimuli from Miranda and Ullman (2007, see §4). Here we only tested the correct/out-of-key contrasts. Target verbs that constituted either correct or violation continuations were time-locked to the target notes in the melodies, and complete sentences were constructed around these frames so that words appeared with all notes in the melodies. The result of these manipulations was a 2 × 2 design shown in Fig. 2 (center panel), which we crossed with melody FAMILIARITY to yield the 2 × 2 × 2 design with eight experimental conditions.

Predictions. Music violations should yield M&U’s patterns: (i) RAN and P600 effects for both familiar and unfamiliar melodies; (ii) only familiar melody violations should elicit their N400? response. It is important to note here that M&U’s music-only study involved a timbre-change detection task, so their participants were attending to the melodies in a task-relevant way. In the present study, participants were tasked with acceptability judgments of the visual/linguistic sequences only, as in previous interference studies targeting syntax¹². However, RAN/P600 violations to music have previously been shown in designs like ours where task-relevant attention was directed

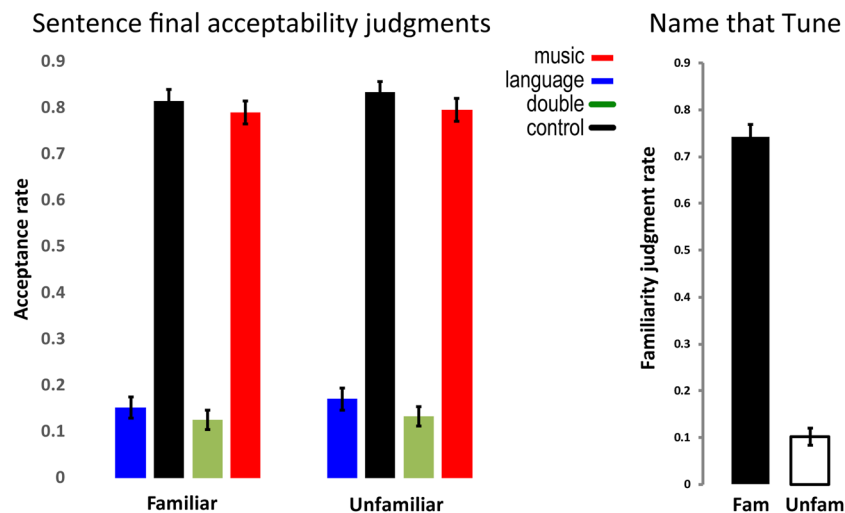


Figure 1. Behavioral responses for end-of-sentence judgment (left) and Name That Tune task (right). Error bars represent 95% CIs.

towards the language stimuli only. Therefore, we can reasonably expect to elicit these responses here, though it is possible that such effects may also be suppressed in situations where music is unattended⁴⁸; (see also Royle *et al.*⁴⁹ for a general discussion of task effects).

Further, we suspected (but did not strongly predict) that the P3a effect previously observed for familiar melodies may not arise here since directing task-relevant attention *away* from the music stimuli and *towards* the complex visual stimuli (sentences) might suppress this response. If so, this design could then also confer the additional possible benefit of allowing a clearer measurement of the music N400 effect to the extent that its distribution may have been partly canceled by the P3a in M&U's data, as discussed above.

We did not have a clear expectation for the music N500. Previous work indicates that this effect may not consistently arise⁵⁰ or may be reduced in amplitude⁴⁸ when task demands direct attention away from the musical/auditory stimuli. However, should we succeed in eliciting an N500, given that there is an indication from one prior interference study⁵⁰ that semantic manipulations in language may interact with this response, this was an anticipated possible outcome in the present study as well (though, as we show below, no such interactions obtained).

We also predicted the linguistic violations would elicit reliable N400s, possibly followed by later anterior negativities and/or P600s. These two post-N400 effects have been occasionally reported elsewhere in the literature for conceptual semantic violations^{51–53}. However, given that such effects are less widely reported and not well-understood, we did not have any prediction regarding whether they would emerge in the present study (see Steinhauer *et al.*⁵¹ for discussion).

Finally, the important open question to be addressed here is whether language/music interaction effects would obtain (CORRECTNESS x KEY interactions). Previous work has suggested that music syntactic violations should not influence language N400s (e.g. Koelsch *et al.*¹² and Steinbeis & Koelsch⁵⁰), so we expected that at least in our unfamiliar melody conditions that finding would replicate (where only main effects of CORRECTNESS and KEY are expected, with no interactions in the N400 latency ranges).

However, it is an open question whether (CORRECTNESS x KEY) interaction effects obtain when N400 effects are elicited in both domains simultaneously, as could be the case for the familiar melody conditions. In fact, previous attempts to carry out experiments much like M&U's or ours found neither music N400 effects^{39,45,46} nor language/music interactions⁵⁴. Thus another possible outcome here would be patterns consistent with this earlier line of work, which is to be expected if the processing of melody familiarity does not in fact engage any systems in common with those underlying language N400s.

Results

Sentence final acceptability & melody familiarity. End of sentence judgment responses (Fig. 1A) yielded main effects of CORRECTNESS [$F(1,35) = 369.45, p < 0.0001$] and KEY [$F(1,35) = 14.05, p = 0.0006$], but no effects of FAMILIARITY and no interactions. Thus, music violations reduced acceptance rates for both correct sentences (interference) and incorrect sentences (facilitation). *Name That Tune* task responses indicated familiar melodies were, in general, known to our participants (Fig. 1B), yielding highly significant differences between melody types [$F(1,35) = 741.91, p < 0.0001$].

Event-related brain potentials. Grand average ERPs in Fig. 2 show all comparisons against matched control items, separating music-only violations (top panels) from language-only and double violations (bottom panels).

Visual inspection of the ERPs in Fig. 2 demonstrates expected early evoked responses (superimposed auditory and visual N1/P2), followed by anterior negativities, prominent N400 responses, and P600 effects for the

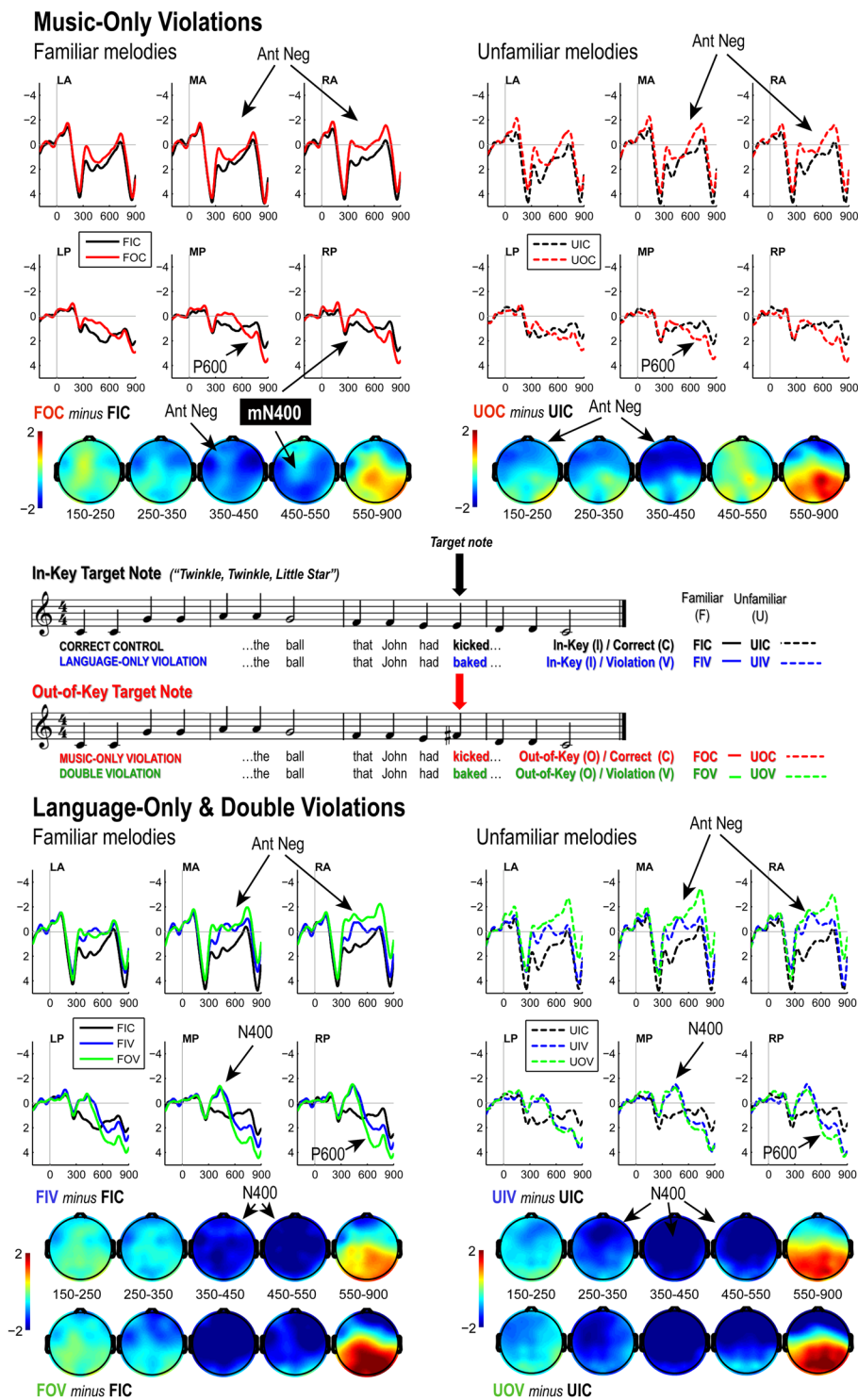


Figure 2. Grand average ERPs (all conditions) and scalp difference maps (Violation minus Control). Music violations shown top panel; Language and double violations in bottom panel. Familiar (lefthand plots) and Unfamiliar (righthand) melody conditions plotted separately. Note that *independent* music and language N400 responses would be expected to yield a larger such effect for the familiar melody double-violation condition (solid green, bottom left) due to additive combination with the music N400. Instead, a sub-additive pattern obtains, indicated by the very similar N400s for language violations whether concurrent music violations obtained (green) or not (blue).

violations compared to control conditions. Crucially, we succeeded in eliciting N400 effects for the language violations (bottom panels of Fig. 2) as well as the expected N400-like negativity for familiar melody **KEY** violations (Fig. 2 top left). In addition, both music and language violations elicited later anterior negativities and P600s.

Observe that the N400 responses in the bottom panels in Fig. 2 are remarkably similar for familiar and unfamiliar melody conditions: both show apparent equivalent N400 responses for language only and double-violation conditions. However, if the apparent music N400 effect for familiar melody music only violations (top left panel) corresponded to independent underlying neural generators, the observed pattern is *not* what we would expect. Instead, what should be seen in that case would be an additive pattern yielding a larger N400 for the familiar double violation case compared to the language only N400 (i.e., the music only and language only N400 responses should sum – *but they do not*). That is, the uniformity of N400 responses in the bottom panels of Fig. 2 clearly reveal a sub-additive N400 response for familiar double violations. This should (and does) yield a significant CORRECTNESS x KEY interaction in our statistical analysis (see below).

The foregoing patterns were confirmed by a top-down analysis of variance of mean amplitude effects for successive time-windows (following M&U²⁴), probing for effects and interactions of KEY and CORRECTNESS, and interactions of these condition factors with scalp location (see Methods). We present these analyses for familiar and unfamiliar melodies in four parts, showing early RAN effects (150–350 ms), linguistic N400s (350–450 ms), our replication of the music N400 response and the demonstration of the crucial CORRECTNESS x KEY interactions (450–550 ms), and late posterior positivities and anterior negativities (550–900 ms). Finally, supplementary analyses were conducted to probe language N400 onset latencies (see Methods), which turned out to vary as a function of melody (un)familiarity (see below). Additional analyses including the factor FAMILIARITY are included where relevant, in particular in demonstrating the presence of the music N400 response for familiar melodies, and its absence in unfamiliar ones.

150–250 & 250–350 ms. Familiar melodies showed no significant effects or interactions involving either KEY or CORRECTNESS in these first two time-windows. There was an indication of the onset of an anterior negativity for familiar melody violations in the 250–350 ms time-window, as can be seen in scalp difference maps (Fig. 2, top left), but the effect of KEY was not significant, even when analyses were restricted to anterior regions of interest (ROIs, *Methods*) [$F(1,35) = 2.85, p = 0.1004$].

In contrast to the familiar melodies, within both these two initial time-windows unfamiliar melodies demonstrated KEY x LATERALITY (LAT) interactions [150–250 ms, $F(2,70) = 5.03, p = 0.0200$; 250–350 ms, $F(2,70) = 3.78, p = 0.0432$] and interactions involving KEY, CORRECTNESS (COR) and both ANTERIOR/POSTERIOR (AP) and LATERALITY [150–250 ms, KEY x COR x AP x LAT: $F(2,70) = 3.31, p = 0.0451$; 250–350 ms, KEY x COR x AP: $F(1,35) = 5.70, p = 0.0225$].

Decomposition of these interactions by ROIs showed that when sentences were Correct, unfamiliar melodies elicited a broadly distributed anterior negativity from 150–250 ms yielding a main effect of KEY over all anterior ROIs [$F(1,35) = 5.29, p = 0.0275$]. However, while significant effects of KEY obtained for the left anterior ROI [$F(1,35) = 9.63, p = 0.0038$], the effect KEY was borderline for the mid-anterior [$F(1,35) = 4.07, p = 0.0514$] and right-anterior ROIs [$F(1,35) = 2.89, p = 0.0978$]. Posterior ROIs showed no effects or interactions involving KEY when sentences were Correct [F 's < 1]. Also, corresponding analyses restricted to sentences with language violations revealed no effects of KEY either [F 's < 1].

From 250–350 ms, analyses within the anterior and posterior ROIs separately showed the same general pattern. When sentences were Correct, there was a significant effect of KEY over anterior ROIs [$F(1,35) = 4.88, p = 0.0337$] that was absent when sentences contained violations [$F < 1$]. There were no effects of KEY for the posterior ROIs [F 's < 1] regardless of presence/absence of language violations.

In addition, in this same 250–350 ms time-window, language violations paired with unfamiliar melodies yielded a main effect of CORRECTNESS [$F(1,35) = 13.07, p = 0.0009$], due the emergence of an N400 effect (Fig. 2, bottom right).

Thus, in these first two time-windows, unfamiliar melody conditions revealed both language and music violation effects, as well as an unanticipated language/music interaction in the form of an apparent suppression of the RAN effect in the double-violation case. During this same period, familiar melody violations showed neither language nor music violation effects.

350–450 ms. Here, robust effects of KEY and CORRECTNESS were evident for both melody types. The language N400 effect that began in the prior time-window for the unfamiliar melody conditions reached its peak, yielding a main effect of CORRECTNESS [$F(1,35) = 35.27, p < 0.0001$]. Similarly, the anterior negativity that was already present for unfamiliar melodies between 150–350 ms became larger, producing a KEY x AP interaction [$F(1,35) = 6.44, p = 0.0158$]. Follow-up analysis confirmed the continued anterior distribution of this negativity for unfamiliar music violations as an effect of KEY was evident over the anterior ROIs [$F(1,35) = 12.15, p = 0.0013$] but only weakly over posterior ROIs [$F(1,35) = 3.46, p = 0.0712$].

Like the unfamiliar melody conditions, familiar melodies also yielded a main effect of CORRECTNESS [$F(1,35) = 22.49, p < 0.0001$] as a result of the emergence of the language N400 effect. Given the difference in the onset latency of language N400 effects as a function of melody familiarity (present already in the previous 250–350 ms window for unfamiliar, but not familiar melodies), we examined (50%) fractional peak latency, peak latency, and peak amplitude measures using a jackknife procedure (see *Methods*), comparing N400 difference waves (Violation – Correct) for familiar versus unfamiliar melodies, both with and without concurrent music violations. These analyses revealed no significant differences in peak latency or amplitude for language N400 effects, but fractional peak latency measures revealed a significant main effect of FAMILIARITY [adjusted $F(1,35) = 5.68, p = 0.0227$], with unfamiliar melodies showing an onset ~50 ms earlier on average (compare scalp difference maps Fig. 2, bottom panels, and right hand plot of language violation effect difference waves in Fig. 3). There was no effect of the presence/absence of music violations on this N400 onset delay.

The familiar melody music violation effect in this time window was differently distributed on the scalp – while the KEY x AP interaction evident for unfamiliar melodies was not significant for familiar ones [$F < 1$], there was

Difference waves (*violation minus control*)

Violation effects + predicted additive profile

Language N400 responses

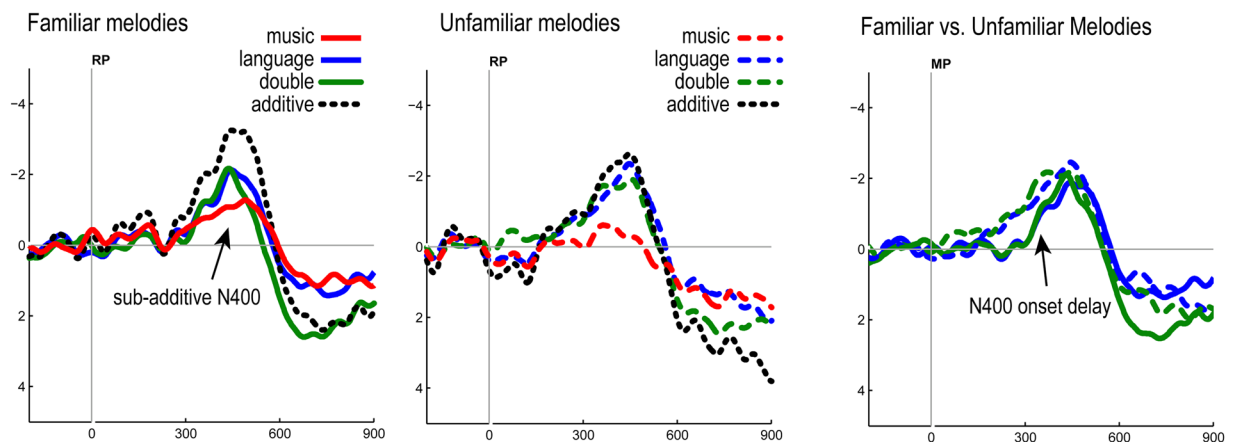


Figure 3. Difference waves for all violations for familiar (left) and unfamiliar (center) melody conditions plotted against predicted additive response profile for double violation conditions; N400 responses for language and double violations superimposed for familiar and unfamiliar melody conditions (right). Note the double-violation (green trace) is sub-additive – corresponding to the Correctness x Key interaction (see also Fig. 2, and main text report of 450–550 ms time-window findings).

a KEY x LATERALITY interaction [$F(1,35) = 3.45, p = 0.0374$]. Follow-ups within the levels of LATERALITY were consistent with a somewhat right lateralized effect of KEY, which was weakly present over the left lateral ROIs [$F(1,35) = 3.92, p = 0.0558$] but significant over both the midline [$F(1,35) = 4.71, p = 0.0369$] and right lateral ROIs [$F(1,35) = 9.42, p = 0.0041$].

450–550 ms. Up to this point, melody conditions were distinguishable only in virtue of the onset timing of both the music and language violation effects (earlier for unfamiliar melody conditions). In the 450–550 ms window, unfamiliar melody conditions showed only language effects [CORRECTNESS: $F(1,35) = 31.13, p < 0.0001$; COR x AP: $F(1,35) = 5.99, p = 0.0196$; COR x AP x LAT: $F(2,70) = 3.31, p = 0.0453$], and no KEY effects or interactions.

In contrast, familiar melody conditions demonstrated both CORRECTNESS effects [$F(1,35) = 16.40, p = 0.0003$] as well as interactions of CORRECTNESS and KEY [$F(1,350) = 6.09, p = 0.0186$] and CORRECTNESS x KEY x AP [$F(2,70) = 4.17, p = 0.0486$]. Follow-up analyses over anterior and posterior ROIs separately showed only a main effect of CORRECTNESS over anterior ROIs [$F(1,35) = 22.04, p < 0.0001$], while posterior ROIs yielded both a CORRECTNESS effect [$F(1,35) = 6.04, p = 0.0191$] and a significant CORRECTNESS x KEY interaction [$F(1,35) = 11.67, p = 0.0016$]. As can be seen in Fig. 2 (top left panel), this interaction was driven by the posterior negativity for familiar melody violations when sentences were correct and the fact that *this effect did not additively combine with the linguistic N400* – note the similar profiles of N400 responses for language-only and double violations in Fig. 2, bottom right panel). Examining this pattern within the levels of CORRECTNESS showed this negativity to be significant over posterior ROIs when sentences were Correct [$F(1,35) = 6.45, p = 0.0157$]; but when language violations occurred, the music violation response trended in the opposite direction, with the violation slightly more positive-going, but not reaching significance [$F(1,35) = 2.78, p = 0.1046$]. A further comparison across the un/familiar melody conditions when sentences were Correct showed a FAMILIARITY x KEY interaction [$F(1,35) = 4.82, p = 0.0348$], confirming the presence of this posterior negativity was unique to the familiar melody music violations.

This could suggest *either* that the music N400 response may have been completely attenuated in double-violation cases or, alternatively, that competition across the processing streams reduced both N400 responses. However, we believe the statistical analysis licenses the former interpretation (complete attenuation of the music N400 in double violations), based on the following reasoning. **First**, the music N400 demonstrates a more circumscribed posterior distribution than the language N400 effect in single violation cases, as can be seen in the voltage maps in Fig. 2, which replicates the distribution reported in Miranda & Ullman's study. **Second**, given this fact, if it were the case that double-violations simultaneously reduced amplitude of *both* the music and language N400 responses to yield the observed sub-additive profile, this ought to be evident more broadly over the scalp in reduced language N400 responses over anterior ROIs as well (given its generally broader distribution).

However, this is not the pattern we see in the data, as discussed above: here we find a three way CORRECTNESS x KEY x A/P interaction that corresponded to a CORRECTNESS x KEY interaction *only over posterior ROIs* – anterior ROIs in contrast showed only a main effect of CORRECTNESS. This strongly suggests that the sub-additive effect was not due to a downward modulation of contributions from *both* language and music processing to the observed double violation N400 amplitude, but rather that the concurrent language violation completely

attenuated the music N400 response. This is also consistent with the observed similarity of the language N400 responses in single and double violations conditions.

To summarize the important findings in this 450–550 ms latency range: the CORRECTNESS x KEY interaction in this time-window reveals a sub-additive N400 response for familiar double violations whereas no such interaction manifested for unfamiliar melodies. In Fig. 3 (left and middle plots) N400 effect difference waves for music, language, and double-violations are shown for both familiar and unfamiliar melodies, plotted against the predicted additive response profile. Again, if the language and music N400 responses were independent, we would expect an additive profile: there would be no CORRECTNESS x KEY interaction for familiar melodies, and the sum of the language & music difference waves (indicated by dotted black curve in Fig. 3) would not differ from the actual double-violation case, contra to fact. As argued above, the overall patterns are consistent with the interpretation that this state of affairs corresponds to a complete attenuation of the music N400 in double-violation cases.

550–900 ms. Finally, in this late latency range both language and music violations yielded anterior negativities and posterior positivities as evident in the robust KEY x AP [$F(1,35) = 26.67, p < 0.0001$] and CORRECTNESS x AP interactions [$F(1,35) = 37.82, p < 0.0001$] between 550–900 ms. There were no interactions between KEY and CORRECTNESS involving these later effects. However, melody FAMILIARITY did modulate the topography of the KEY effects [FAMILIARITY x KEY x AP: $F(1,35) = 4.13, p = 0.0497$; FAMILIARITY x KEY x LATERALITY: $F(2,70) = 4.07, p = 0.0243$]. In particular, like the earlier RAN-type negativities, unfamiliar melody late anterior negativities (N500) showed no indication of lateralization, while familiar melodies were right lateralized. These lateralization differences can be seen most clearly in the double-violation ERP responses in Fig. 2, where these effects were particularly large (due to the additivity of apparently *independent* language and music late anterior negativities).

General summary. The results of our experiment can be summarized as follows. **First**, both music only and language only violations replicated previously reported patterns. Violations of KEY for both melody types elicited RANs followed by later N500 and P600 effects, while only familiar melodies yielded the additional posterior negativity (N400) which emerged following (and slightly overlapping with) the RAN, just as reported in M&U²⁴ (though differing from M&U²⁴ in onset timing, see below). The conceptual semantic violations in the visually presented linguistic stimuli also replicated previously reported patterns, including robust N400 effects followed by P600s and anterior negativities (see Steinhauer *et al.*⁵¹). Two findings from the original music only study did not replicate here, namely the anterior P3a and the larger P600 for familiar melody violations. The absence of these previously reported effects, as we discuss below, is relevant to our preliminary interpretation of the present data.

Second, in terms of cross-domain interference effects, **four** findings (I–IV) are relevant:

- (I) Onset delay of the language N400. Though language N400 effects did not show amplitude modulations tied to concurrent music processing, onset timing was impacted by (un)familiarity, independent of presence/absence of music violations. The language N400 demonstrated an onset delay when the corresponding melody was familiar (Fig. 3, right panel).
- (II) Onset delay of RAN/N400 complex for familiar melodies. Familiar music only violations also demonstrated an onset delay. In the previous study testing these melodies in isolation²⁴, familiar and unfamiliar conditions elicited RANs with indistinguishable onset timing, with the familiar melody N400 effect following (and slightly overlapping) the RAN. Here, though the RAN and N400 effects for familiar melodies occurred together in sequential manner closely resembling the previous report, this complex of responses did not arise until 350–550 ms, whereas the RAN for *unfamiliar* melodies was already detectable in the 150–250 ms latency range.
- (III) Sub-additive N400 response in familiar melody double-violation conditions. If music and language N400 responses were independent, we would expect to see an additive response profile. For familiar melodies, additivity would present with double violation N400s, expected to be larger than the language only N400 response, with no statistical interaction (i.e., just KEY and CORRECTNESS main effects). However, because the unfamiliar melodies should not elicit the music N400 response, in the unfamiliar case we would expect to see similar if not identical N400 responses whether there was a concurrent music violation or not. While this is exactly what we see for unfamiliar melodies, the music N400 effect for familiar melodies appears to have been suppressed in double violations, yielding indistinguishable N400 amplitudes for familiar and unfamiliar melody condition language-only and double violation cases. In the familiar melody case, this corresponds to the crucial KEY x CORRECTNESS interaction that was one of our anticipated possible findings.
- (IV) Unfamiliar melody RAN effect was delayed in double violations. Presence of linguistic semantic violations suppressed/delayed the RAN for unfamiliar melodies. This was not an anticipated outcome of the present study, though as we discuss below it may not be without precedent either (e.g., see Loui *et al.*⁴⁸).

Discussion

The present data make a strong case that the posterior negativity for familiar melody violations documented in Miranda and Ullman (2007) is best viewed as a genuine N400 response. Both familiar and unfamiliar melody conditions here replicated music-syntactic violation responses (RAN and later N500/P600 effects), and only familiar melody violations yielded an additional N400-like negativity directly following the RAN.

Nailing down the functional significance of this component is an important finding on its own. We will henceforth refer to this effect as the **music N400 (mN400)**, as our data warrant removal of the ?-mark from M&U's

“N400?”. Concurrent auditory processing of the familiar melodies that elicit this mN400 response delayed the onset of the language N400 effect independent of the presence/absence of music violations (see Fig. 3, right panel, and scalp voltage maps for language N400 responses in Fig. 2). This N400 onset delay was seen relative to the unfamiliar melody case and in absence of any other modulations of the N400 in the unfamiliar conditions. This pattern serves to replicate previous reports that music syntactic violations *on their own* do not influence language N400 effects^{12,20}. We take this dissociation to indicate competition for resources in the familiar melody conditions only, involving the management of the activation of language and music LTM semantic memory systems.

Importantly, the language N400 onset delay for familiar melody conditions is not the only relevant finding in these data. When sentences were well-formed, the RAN/mN400 complex for violations in familiar melodies was delayed. In Miranda and Ullman (2007), familiar and unfamiliar melody RAN effects showed equivalent onsets (~150–270 ms) with the familiar melody mN400 onset only slightly later (220–380 ms). Here, the familiar melody RAN did not emerge until 350–450 ms with the mN400 arising at 450–550 ms. That is, familiar melody violation responses were delayed by concurrent language processing in the present study by as much as ~200 ms (judging by comparison to the previously reported data). It may be no accident that the magnitude of this delay in familiar melody violation responses roughly corresponds with estimates of the timing of early aspects of lexical access^{55,56}.

Finally, the fact that our familiar melody double-violations yielded a sub-additive N400 response is also strongly consistent with M&U’s interpretation of their posterior negativity as a *bona fide* N400 effect. Further, as noted above (see 450–550 ms time-window report in Results), the pattern of effects suggests that this sub-additive response is best understood as a complete suppression of the music N400 response in double-violation cases. As can be seen in the data visualizations (Fig. 2), and as our statistical analysis showed, the language N400 response appears completely unperturbed by the concurrent music violation (note the similarity of average ERP curves for language only (blue traces) and double violation cases (green traces) across the scalp for familiar melodies in Fig. 2, and the overall similarity to the corresponding effects for the unfamiliar melodies).

Before we turn to other effects we have documented here, and possible ways to interpret these patterns overall, it is important to note that the present findings concerning the N400 and language/music processing differ in important ways from previous findings concerning language/music syntactic processing and the corresponding LAN/RAN response profiles. Whereas previous work¹² showed a LAN (language effect) attenuation in double violation cases, here we show different kinds of cross-domain influences in both directions (e.g., familiarity of melodies delayed language N400 onset; double violations attenuated music N400 amplitude; concurrent language processing delayed the RAN/mN400 onset, etc.).

In addition, another important pattern documented here involving an influence of concurrent language processing on music is the early sub-additive response over anterior scalp regions for the *unfamiliar* melody double-violations, which appears to be best described as indicating an early suppression/delay of the RAN. This seems important to consider relative to the fact, just mentioned above, that *both* the RAN and mN400 effects were delayed in the familiar melody condition when sentences were correct (compared to M&U’s music-only results), and that the RAN was generally later in onset for the familiar cases (see scalp voltage maps in Fig. 2). The only early onset RAN effect we find here was for single (music) violations when melodies were unfamiliar. How can we interpret these patterns?

Since the language violations we tested here turned on lexical/conceptual semantic properties of target words, the fact that these violations would influence the RAN effect for unfamiliar melodies could be seen as an unexpected result based on previous findings^{12,20}. Similar considerations arise for the delay of the RAN for familiar melody violations that arose independent of sentence correctness. Simultaneous demands on LTM access/retrieval across processing streams could be appealed to in order to explain the delay of the mN400 (and the language N400), but why should the RAN effects also be delayed? And why should this RAN delay occur *uniformly* for familiar melodies with no apparent influence of language Correctness, but show an interaction profile for the unfamiliar conditions with RAN delay only when both streams were simultaneously violated?

Given these patterns, one might be tempted to turn to accounts that appeal to demands on general attention¹⁰. However, such an explanation would not easily account for the full range of findings in the present study (in addition to other worries that can be raised about such accounts, see Slevc & Okada⁷ and Kunert *et al.*⁵⁷). It is possible that for unfamiliar melodies, when sentences were well-formed, there were only weak demands on attentional resources and so the music RAN effect showed an onset timing consistent with the previous music-only study. In contrast, when semantic violations were present, this could have increased the processing load, attenuating the task irrelevant auditory channel, thus suppressing/delaying the RAN. It has already been demonstrated elsewhere⁴⁸ that RAN effects can be suppressed when attention is directed away from music, so this part is plausible.

But we might then ask if this reasoning can explain why the RAN effect was delayed for *familiar* melodies when there was *no* concurrent language violation. It is possible that familiar melodies generally capture less attentional resources in virtue of their *predictability*. It is known, for example, that attention and predictability can have opposite effects on early evoked responses (e.g., smaller N1 effects for predicted items, larger N1 effects for attended items⁵⁸). The predictability of familiar melodies may thus cause less attention to be given to the auditory stream, mirroring the previous argument of attenuation of the auditory channel by semantic violations in unfamiliar melodies (and creating a similar delay in the RAN effect). But if familiar melodies were less well-attended in virtue of their predictability, why would their concurrent processing cause a delay in the language N400 effects? This latter aspect of the response pattern suggests that attention was directed *away* from the visual/language processing stream by the presence of familiar melodies (i.e., that they demanded *more*, not *less*, attention).

Likewise, several other aspects of our ERP data, relative to patterns in Miranda and Ullman (2007), seem to weigh against any *simple* appeal to attention here. First, that our task directed attention away from the music stimuli is consistent with the absence of the P3a effect for familiar melodies found in the previous music-only study *and* with the fact that familiar and unfamiliar melodies elicited equivalent P600 effects (whereas previous findings showed larger P600s for familiar melody violations^{24,39}). Reduction of P600s as a function of presence/

absence of task-relevant attention is expected based on other prior work (see review⁴⁹). Thus, asymmetries in the response profiles, plausibly linked to attention differences across familiar and unfamiliar melodies in the music-only study²⁴, could be said to have been neutralized here where the task directed attention to the visual/language processing stream.

This neutralization of likely markers of attentional asymmetries across melody types may also have very early correlates in the ERP record. An undiscussed pattern in M&U's data (Fig 3 & 4 in M&U²⁴), pertains to the early auditory P1-N1-P2 complex, where P2 responses were noticeably larger for unfamiliar compared to familiar melodies. In fact, inspection of their data suggests amplitude differences as large as $\sim 2 \mu\text{V}$ s related to (un)familiarity, though these apparent differences were not analyzed. This may be thought of in terms of the attention/prediction trade-off mentioned above, with the predictability of familiar melody continuations resulting in lower demands on early perceptual level processing.

Here, in contrast, auditory and visual N1-P2 effects superimposed, yielding larger amplitudes in general but with no indication of larger P2 amplitudes when melodies were unfamiliar (Fig. 2). Like the absence of the P3a and the familiar/unfamiliar P600 amplitude differences, these observations are consistent with the absence of general attentional asymmetries between familiar and unfamiliar melodies in the current study that were otherwise present when the melodies were processed in isolation and attended to in a task-relevant way.

What must be accounted for here is the *symmetry* of interference for familiar melodies, and the *asymmetry* of the interference when melodies were unfamiliar. It may be that an account based in considerations involving attention can be reconciled with views positing shared syntactic processing resources within models that view working memory as attention-regulated activation of long term memory (LTM) representations (see LaRocque *et al.*⁵⁹ for overview).

In the unfamiliar melody cases, since detection of the music violations does not implicate activation of specific long term memory representation (i.e., of the “musical lexicon”, in the sense similar to Peretz *et al.*⁶⁰), there is no conflict with the sub-systems at work in language processing involving prediction and access/retrieval of lexical/conceptual semantic information associated with incoming words when sentences were well-formed. Accordingly, RAN effects for unfamiliar melodies are registered with a timing similar to when music is processed in isolation, so long as there are no disruptions in the language processing stream. In contrast, for familiar melodies, attentional control mechanisms of LTM activation must necessarily already be subjected to simultaneous demands across the language and music streams even prior to encounters with target notes/words in our paradigm.

Development of this line of explanation seems capable of accounting for the bi-directional effects on the N400 for language and music in the familiar melody conditions. If correct, this would also suggest that the RAN for familiar melodies is not automatically triggered by the deviant out-of-key notes independently of the processes involved in access/retrieval of corresponding musical LTM representations, at least in circumstances when music is not attended to. From this perspective it would be important to determine what effect a task directing attention to the auditory/music stream would have. In principle, this could result in bringing the onset latency of familiar and unfamiliar melody RAN effects back into their previously reported alignment. An open question would be whether this would also result in a similarly earlier onset for the mN400, more consistent with the timing reported in the previous music-only study, or not (we are investigating these questions directly in the context of our larger scale study).

If RAN effects are generally sensitive to attention in this way (consistent with Loui *et al.*⁴⁸), then the suppression/delay of the unfamiliar melody RAN when there was a concurrent language violation could be given the simple treatment sketched above. Given the lack of competition between the streams in terms of attentional control of the activation of LTM representations, presence of a language violation captures attentional resources unhindered by demands from the music processing stream, resulting in the delay of the RAN in double-violations, but no influences in the other direction (i.e., no effects of unfamiliar melody processing on language). In the familiar melody double-violation conditions we expect a different pattern, as both streams involve simultaneous disruptions of activity underlying the N400. In this case, having attentional resources swamped by processing of the language violation could be expected to suppress the mN400 response – precisely the pattern evident in our data.

Of course, this line of interpretation requires further fleshing out, and several follow-up inquiries are required to demonstrate how robust (replicable) and how general these N400 interference effects may be (see below). Our point here is simply that an approach that takes attention into account in formulating the nature of cross-domain effects concerning working/long-term memory and language/music relationships seems to be recommended. That is, even if we concur with arguments that *general* appeals to attention are *insufficient* on their own to provide an account of cross-domain interference effects^{7,57}, it may be that appeals to attention are nonetheless *necessary* for any adequate account of such patterns.

In addition to the open questions discussed above, as mentioned in §1 many but not all of the familiar melodies we adopted from the M&U study have associated lyrics. Thus, it will be important to determine whether the presence of the mN400 (and/or the cross-domain interaction patterns) depend on this feature or not. Previous work has provided documentation of both dissociations between melody and lyric processing^{54,60} as well as their tight linkage⁶¹. However, it is important to point out that this issue does not diminish the interest and importance of the present findings. Indeed, showing as we have here for the first time that N400s can indeed be elicited in connection with isolated music stimuli, and can in addition yield interference profiles in paradigms like the one employed here, opens up a range of possibilities regarding future follow-up work on the nature of musical LTM, which is of course multi-dimensional (linking to semantic, episodic, and emotional/affective memory systems, in addition to the systems supporting LTM of lyrics). This finding should be significant also for studies of musical memory in special populations (e.g., in patient groups, examinations of age-related cognitive changes, etc.^{62–66}). Further, additional questions arise about *how general* this type of cross-domain N400 interaction may be⁶⁷, and what impact the potential generality of such interactions would have on our understanding of musical

semantics^{42,68}. For example, N400-like effects have also been reported for familiar versus unfamiliar environmental sounds⁶⁹ and it is an open question as to whether such effects would cause interference with language (or music) N400 responses as well (see also work examining N400 responses to words versus environmental noises^{70,71}).

Finally, the fact that it is possible to elicit such N400 interference profiles across these cognitive domains (and modalities) suggests possible opportunities for leveraging these findings in efforts to adjudicate between ongoing disputes concerning the etiology of this ERP response (access/retrieval? binding? integration? inhibition?), and how it relates to general concerns about working memory and attention (see also discussion in Kutas & Federmeier²⁶ regarding the N400 and attention). Our findings also recommend connecting this type of inquiry with work on unimodal versus cross-modal attention⁷² and multi-sensory integration more broadly⁷³.

Methods

Participants. Data reported reflects $N = 36$ (20 M; 16 F; age 18–31; $M = 21.08$, $SD = 2.77$). All were right handed (assessed by modified version of the Edinburgh Handedness Inventory⁷⁴), had normal or corrected to normal vision, normal hearing, and no history of neurological pathology. Music experience ranged from 0–30 years singing or playing an instrument ($M = 7.76$ years; assessed by questionnaire adapted from Wei Looi⁷⁵). Participants' scores ranged from 1.00 to 4.86 on the Innovative Musical Aptitude subscale of the Brief Music Experience Questionnaire⁷⁶. This subscale measures musical performance ability (1–5 low-to-high experience scale). Fourteen participants (~39%) reported they currently play an instrument or sing. All participants gave informed written consent and were given research credit or were paid for participation.

Materials. *Musical stimuli.* These were drawn from M&U²⁴. *Familiar* melodies were chosen from six different categories: traditional/folk, children's music, patriotic songs, holiday songs, classical music, and pop/theme music. Many, although not all, of these melodies have lyrics associated with them. *Unfamiliar* (novel) melodies were composed to be pair-wised matched to familiar melodies. Of the 120 melodies in M&U²⁴, we used 112 (8 excluded due to very fast tempo). Individual note onsets were extracted for use in timing the visual word-by-word sentence presentation. Melodies were grouped together in sets of four, creating twenty-eight groups of melodies (groups had similar total number of notes).

Sentence stimuli. These were constructed around object relative clauses containing target VERBS, created in 4-tuples (e.g., the ball that the boy has KICKED/*BAKED; the pie that the boy has *KICKED/BAKED) to rotate target words across conditions. 4-tuples were duplicated and edited to allow for both inflected and uninflected target verbs (e.g., ... the ball that the boy will KICK...), creating a total of 56 8-tuple groups of sentence frames.

Melody-sentence pairing and list creation. The melody list was duplicated so that each of the 28 groups of melodies contained two sets of its four unique melodies. Each group of melodies was randomly assigned to two 8-tuples of object-relative clause frames, such that each melody was assigned to two matched pairs of correct/violation sentence frames (one from each 8-tuple). Each sentence frame pair was then matched with each of the four versions of that melody. This created 16 unique items for each melody, for a total of 64 unique items for each group of melodies. This was done for all 28 groups of melodies, yielding 1792 unique items (sentence-frame/melody pairings).

Sentence beginnings/endings were added to create a context for each relative clause frame. Critical target verbs were aligned to target notes of the melodies; each syllable of text occurred with one melody note. Target verbs and the immediately preceding word were always monosyllabic.

The master list of sentence/melody pairings was factored into eight presentation lists, each consisting of 224 sentence-melody pairs (28 per condition in the $2 \times 2 \times 2$ (FAMILIARITY \times CORRECTNESS \times KEY design). Items were distributed across lists such that each list contained both a familiar and unfamiliar version of each melody with its matched sentence. In-key and Out-of-key versions of any given melody never appeared in the same list. Unfamiliar versions of each melody always preceded the familiar one. Items were rotated such that no list contained both the correct/violation version of any sentence and were pseudo-randomized.

Procedure. All procedures in the present study were carried out in accordance with relevant guidelines, precisely as specified in the experimental protocol approved by Stony Brook University IRB.

Session 1: EEG data collection. Participants were seated in a booth facing a computer monitor while wearing ear-insert earphones. Sentence-melody pairs were presented with words appearing one at a time in the center of the computer monitor. Each word was presented at note onset with timing determined by the length of the note, until it was replaced by the next word (at onset of following note). Both target and pre-target notes always had a uniform duration of 600 ms. Participants were made aware that they may hear some familiar melodies along with some unfamiliar ones, and that some melodies may include “out of key” or “sour” notes. Participants were instructed, however, to evaluate only the visual/linguistic stimuli, producing an end of sentence judgment for all trials. Sentences were judged as “good” or “bad”, indicated by a mouse click.

Session 2: “Name That Tune”. Participants returned for a second session (always within less than one week from session 1) during which Familiar Correct and Unfamiliar Correct versions of the musical stimuli presented in Session 1 were re-presented, but without sentence stimuli. Melodies were played over speakers and presented in the same order as in Session 1. Participants indicated by mouse click songs they recognized and were asked to say the name of the tune out loud if identified. If they did not know/remember the name of the song, but were able to give other information (lyrics, semantic category, etc), they were instructed to provide this. Familiarity

judgments were recorded automatically, while any comments supplied by the participant were recorded manually by the experimenter.

EEG recording, data processing, and analysis. Scalp EEG was recorded from 32 electrodes plus H/VEOG (512 Hz sampling rate), referenced offline to averaged left/right mastoids (BioSemi Active2 amplifiers; online bandpass of 0.01–125 Hz). Trials contaminated with eyeblinks or other artifacts were excluded automatically. All channels were filtered offline (0.1–30 Hz) and single subject waveforms were averaged over 1200 ms epochs following the target note, including –200 to 0 ms baseline correction interval.

Repeated measures ANOVAs were conducted to analyze ERP data over midline and lateral regions of interest (ROIs) in a 2×3 grid with two levels of ANTERIOR/POSTERIOR (AP) and three levels of LATERALITY (left/mid/right) as follows: Left-Anterior = AF3/F3/F7, Mid-Anterior = FP1/FP2/Fz, Right-Anterior = AF4/F4/F8, Left-Posterior = CP5/P3/PO3, Mid-Posterior = Cz/CP1/CP2/Pz/O1/Oz/O2, Right-Posterior = CP6/P4/PO4. All analyses of mean amplitude measures within given time-windows (150–250, 250–350, 350–450, 450–550, 550–900 ms) were carried out over these ROIs. Analyses included the language factor CORRECTNESS (Correct/Violation) and the music factor KEY (In-key/Out-of-key). These analyses were first carried out for familiar and unfamiliar melodies separately, to probe for predicted effects. Additional analyses including the factor FAMILIARITY were carried out to demonstrate statistical interactions where relevant, in particular concerning the presence/absence of the music N400 response. Greenhouse-Geiser corrections were performed where relevant (corrected *p*-values are reported throughout). Onset timing of language N400 effects were examined, including peak amplitude and both peak latency and (50%) fractional peak latency measures, analyzed using a jackknife method⁷⁷. Reported *F*-values for the jackknife analyses were adjusted by dividing by $(N-1)^2$. Sequence final acceptance rates (% accepted) for judgment task and familiarity judgments (% judged as familiar) in the follow-up *Name That Tune* task, were also examined with repeated measures ANOVA.

Ethics. All participants provided informed written consent. The study was approved by the Stony Brook University IRB.

Data availability

Datasets supporting this report are included as electronic supplementary information (see SI-Overview).

Received: 15 May 2018; Accepted: 3 April 2020;

Published online: 08 July 2020

References

- Patel, A. D. Language, music, syntax and the brain. *Nat Neurosci.* **6**(7), 674–681 (2003).
- Patel, A. D. *Music, language, and the brain*. New York: Oxford University Press (2008).
- Peretz, I., Vuvan, D., Lagrois, M.-É. & Armony, J. L. Neural overlap in processing music and speech. *Philos Trans R Soc B.* **370**(1664), 1–8 (2015).
- Peretz, I. & Coltheart, M. Modularity of music processing. *Nat Neurosci.* **6**(7), 688–691 (2003).
- Besson, M. & Schön, D. Comparison between language and music. *Ann NY Acad Sci.* **930**(1), 232–258 (2001).
- Heffner, C. C. & Slevc, L. R. Prosodic structure as a parallel to musical structure. *Front Psychol.* **6**, 1–14 (2015).
- Slevc, L. R. & Okada, B. M. Processing structure in language and music: A case for shared reliance on cognitive control. *Psychon Bull Rev.* **22**(3), 637–652 (2015).
- Van de Cavey, J. & Hartsuiker, R. J. Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition.* **146**, 172–184 (2015).
- Kljajević, V. Is Syntactic Working Memory Language Specific? *Psihologija.* **43**(1), 85–101 (2010).
- Rogalsky, C., Rong, F., Saberi, K. & Hickok, G. Functional anatomy of language and music perception: Temporal and structural factors investigated using functional Magnetic Resonance Imaging. *J Neurosci.* **31**(10), 3843–3852 (2011).
- Koelsch, S. & Siebel, W. A. Towards a neural basis of music perception. *Trends Cogn Sci.* **9**(12), 578–584 (2005).
- Koelsch, S., Gunter, T. C., Wittfoth, M. & Sammler, D. Interaction between Syntax Processing in Language and in Music: An ERP Study. *J Cogn Neurosci.* **17**(10), 1565–1577 (2005).
- Chomsky, N. *Aspects of a theory of syntax*. MIT Press (1965).
- Lerdahl, F. A. & Jackendoff, R. S. *A generative theory of tonal music*. MIT Press (1983).
- Rohrmeier, M. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music.* **5**(1), 35–53 (2011).
- Rohrmeier, M. & Koelsch, S. Predictive information processing in music cognition. A critical review. *Int J Psychophysiol.* **83**(2), 164–175 (2012).
- Fedorenko, E., Patel, A. D., Casasanto, D., Winawer, J. & Gibson, E. Structural integration in language and music: evidence for a shared system. *Mem Cognit.* **37**(1), 1–9 (2009).
- Slevc, L. R., Rosenberg, J. C. & Patel, A. D. Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon Bull Rev.* **16**(2), 374–381 (2009).
- Hoch, L., Poulin-Charronnat, B. & Tillmann, B. The influence of task-irrelevant music on language processing: Syntactic and semantic structures. *Front Psychol.* **2**, 1–10 (2011).
- Carrus, E., Pearce, M. T. & Bhattacharya, J. Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations. *Cortex.* **49**(8), 2186–2200 (2013).
- Molinaro, N., Barber, H. & Carreiras, M. Grammatical agreement processing in reading: ERP findings and future directions. *Cortex.* **47**(8), 908–930 (2011).
- Steinhauer, K. & Drury, J. E. On the early left anterior negativity (ELAN) in syntax studies. *Brain Lang.* **120**(2), 135–162 (2012).
- Patel, A. D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J. Processing Syntactic Relations in Language and Music: An Event-Related Potential Study. *J Cogn Neurosci.* **10**(6), 717–733 (1998).
- Miranda, R. A. & Ullman, M. T. Double dissociation between rules and memory in music: an event-related potential study. *Neuroimage.* **38**(2), 331–345 (2007).
- Koelsch, S. & Jentschke, S. Differences in electric brain responses to melodies and chords. *J Cogn Neurosci.* **22**(10), 2251–2262 (2009).

26. Kutas, M. & Federmeier, K. D. Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annu Rev Psychol.* **62**, 621–647 (2011).
27. Groussard, M. *et al.* Musical and verbal semantic memory: two distinct networks? *Neuroimage.* **49**(3), 2764–2773 (2010).
28. Satoh, M., Takeda, K., Nagata, K., Shimosegawa, E. & Kuzuhara, S. Positron-emission tomography of brain regions activated by recognition of familiar music. *AJNR Am J Neuroradiol.* **27**(5), 1101–1106 (2006).
29. Plailly, J., Tillmann, B. & Royet, J. P. The feeling of familiarity of music and odors: the same neural signature? *Cereb Cortex.* **17**(11), 2650–2658 (2007).
30. Kutas, M. & Hillyard, S. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science.* **207**(4427), 203–205 (1980).
31. Hagoort, P. Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *J Cogn Neurosci.* **15**(6), 883–899 (2003).
32. Baggio, G. & Hagoort, P. The balance between memory and unification in semantics: a dynamic account of the N400. *Lang Cogn Process.* **26**(9), 1338–1367 (2011).
33. Debrulle, J. B. The N400 potential could index semantic inhibition. *Brain Res Rev.* **56**(2), 472–477 (2007).
34. Cummings, A. *et al.* Auditory semantic networks for words and natural sounds. *Brain Res.* **1115**(1), 92–107 (2006).
35. Koelsch, S. *et al.* Music, language, and meaning: brain signatures of semantic processing. *Nat Neurosci.* **7**(3), 302–307 (2004).
36. Painter, J. G. & Koelsch, S. Can out-of-context musical sounds convey meaning? An ERP study on the processing of meaning in music. *Psychophysiology.* **48**(5), 645–655 (2011).
37. Daltrozzo, J. & Schön, D. Conceptual processing in music as revealed by N400 effects on words and musical targets. *J Cogn Neurosci.* **21**(10), 1882–1892 (2008).
38. Daltrozzo, J. & Schön, D. Is conceptual processing in music automatic? An electrophysiological approach. *Brain Res.* **1270**, 88–94 (2009).
39. Besson, M. & Faïta, F. An event-related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J Exp Psychol Hum Percept Perform.* **21**(6), 1278 (1995).
40. Williamon, A. & Egner, T. Memory structures for encoding and retrieving a piece of music: an ERP investigation. *Cogn Brain Res.* **22**(1), 36–44 (2004).
41. Koelsch, S., Gunter, T., Friederici, A. D. & Schröger, E. Brain Indices of Music Processing: ‘Nonmusicians’ are Musical. *J Cogn Neurosci.* **12**(3), 520–541 (2000).
42. Koelsch, S. Towards a neural basis of processing musical semantics. *Phys Life Rev.* **8**(2), 89–105 (2011).
43. Featherstone, C. R., Morrison, C. M., Waterman, M. G. & MacGregor, L. J. Semantics, syntax, or neither? A case for resolution in the interpretation of N500 and P600 responses to harmonic incongruities. *PLoS ONE.* **8**(11), e76600 (2013).
44. Polich, J. Updating P300: An integrative theory of P3a and P3b. *Clin Neurophysiol.* **118**(10), 2128–2148 (2007).
45. Besson, M. & Macar, F. An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology.* **24**(1), 14–25 (1987).
46. Paller, K. A., McCarthy, G. & Wood, C. C. Event-related potentials elicited by deviant endings to melodies. *Psychophysiology.* **29**(2), 202–206 (1992).
47. Frey, A., Aramaki, M. & Besson, M. Conceptual priming for realistic auditory scenes and for auditory words. *Brain Cogn.* **84**(1), 141–152 (2014).
48. Loui, P., Gent-T, J. T., Torpey, D. & Woldorff, M. Effects of attention on the neural processing of harmonic syntax in Western music. *Cogn Brain Res.* **25**(3), 678–687 (2005).
49. Royle, P., Drury, J. E. & Steinhauer, K. ERPs and task effects in the auditory processing of gender agreement and semantics in French. *Ment Lex.* **8**(2), 216–244 (2013).
50. Steinbeis, N. & Koelsch, S. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb Cortex.* **18**(5), 1169–1178 (2008).
51. Steinhauer, K., Drury, J. E., Portner, P., Walenski, M. & Ullman, M. T. Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia.* **48**(6), 1525–1542 (2010).
52. Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. & Ullman, M. T. The influence of language proficiency on lexical semantic processing in native and late learners of English. *J Cogn Neurosci.* **24**(5), 1205–1223 (2012).
53. Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, A. Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Mem Cognit.* **23**(4), 477–494 (1995).
54. Besson, M., Faïta, F., Peretz, I., Bonnel, A.-M. & Requin, J. Singing in the brain: independence of lyrics and tunes. *Psych Sci.* **9**(6), 494–498 (1998).
55. Brennan, J., Legnos, C., Embick, D. & Roberts, T. P. Spectro-temporal correlates of lexical access during auditory lexical decision. *Brain Lang.* **133**, 39–46 (2014).
56. Mellem, M. S., Friedman, R. B. & Medvedev, A. V. Gamma- and theta-band synchronization during semantic priming reflect local and long-range lexical-semantic networks. *Brain Lang.* **127**(3), 440–451 (2013).
57. Kunert, R., Willems, R. M. & Hagoort, P. Language influences music harmony perception: effects of shared syntactic integration resources beyond attention. *R. Soc. open sci.* **3**(2), 150685 (2016).
58. Lange, K. The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous finding on the auditory N1 with opposite effects of attention and prediction. *Front Hum Neurosci.* **7**, 1–14 (2013).
59. LaRocque, J. J., Lewis-Peacock, J. A. & Postle, B. R. Multiple neural states of representation in short-term memory? It’s a matter of attention. *Front Hum Neurosci.* **8**, 5 (2014).
60. Peretz, I. *et al.* Functional dissociations following bilateral lesions of auditory cortex. *Brain.* **117**(6), 1283–1301 (1994).
61. Peretz, I., Radeau, M. & Arguin, M. Two-way interactions between music and language: Evidence from priming recognition of tune and lyrics in familiar songs. *Mem Cognit.* **32**(1), 142–152 (2004).
62. Schulkind, M. D. Conceptual and perceptual information both influence melody identification. *Mem. Cognit.* **32**, 841–851, <https://doi.org/10.3758/BF03195873> (2004).
63. Hailstone, J. C., Omar, R. & Warren, J. D. Relatively preserved knowledge of music in semantic dementia. *J. Neurol. Neurosurg. Psychiatry.* **80**, 808–809, <https://doi.org/10.1136/jnnp.2008.153130> (2009).
64. Weinstein, J. *et al.* Preserved musical semantic memory in semantic dementia. *Arch. Neurol.* **68**, 248–250, <https://doi.org/10.1001/archneurol.2010.364> (2011).
65. Vanstone, A. D. *et al.* Episodic and semantic memory for melodies in Alzheimer’s disease. *Music Percept.* **29**, 501–507, <https://doi.org/10.1525/mp.2012.29.5.501> (2012).
66. Sikka, R., Cuddy, L. L., Johnsrude, I. S. & Vanstone, A. D. An fMRI comparison of neural activity associated with recognition of familiar melodies in younger and older adults. *Front. Neurosci.* **9**, 1–10, <https://doi.org/10.3389/fnins.2015.00356> (2015).
67. Aramaki, M., Marie, C., Kronland-Martin, R., Ystad, S. & Besson, M. Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *J Cogn Neuro.* **22**, 2555–2569 (2009).
68. Besson, M., Frey, A. & Aramaki, M. Is the distinction between intra- and extra-musical meaning implemented in the brain? Comment on “Towards a neural basis of processing musical semantics” by Stefan Koelsch. *Physics of Life Reviews.* **8**, 112–113 (2011).
69. Kirmse, U., Jacobsen, T. & Schroger, E. Familiarity affects environment sound processing outside focus of attention: An event-related potential study. *Clin Neurophysiol.* **120**(5), 887–896 (2009).

70. Hendrickson, K., Friend, M., Walenski, M. & Love, T. The organization of words and environmental sounds in memory. *Neuropsychologia* **69**, 67–76 (2015).
71. Hendrickson, K., Love, T., Walenski, M., & Friend, M. The organization of words and environmental sounds in the 2nd year: Behavioral and electrophysiological evidence. *Developmental Science*, e12746 (2018).
72. Shandu, R. & Dyson, B. J. Cross-modal perceptual load: the impact of modality and individual differences. *Exp Brain Res* **234**(5), 1279–1291 (2016).
73. Cappe, C., Thut, G., Romei, V. & Murray, M. M. Auditory-visual multi-sensory interactions in humans: timing topography, directionality, and sources. *J Neurosci.* **30**(38), 12572–12580 (2010).
74. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. **9**(1), 97–113 (1971).
75. Wei Looi, V. MLEQ and MTEQ in Music Perception of Cochlear Implant Users. (Doctoral Dissertation). Retrieved December 2012, from http://ir.canterbury.ac.nz/bitstream/10092/2200/1/12608302_thesis%20final.pdf (2006).
76. Werner, P. D., Swope, A. J. & Heide, F. J. The Music Experience Questionnaire: development and correlates. *J Psychol.* **140**(4), 329–345 (2006).
77. Miller, J., Ulrich, R. & Schwarz, W. Why jackknifing yields good latency estimates. *Psychophysiology*. **46**(2), 300–312 (2009).

Acknowledgements

We would like to warmly acknowledge Drs. R. Miranda and M. T. Ullman for making their music stimuli available to us for the use in the present study. We would also like to thank the Experimental Linguistics lab at Stony Brook University, including the many Research Assistants who assisted in data collection. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

J.E.D. conceived and designed the experiments, supervised collection and analyses of the data as well as manuscript preparation. N.C.R. created all language materials, collected and analyzed the data, and wrote large parts of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to N.C.-R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020