Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception

Nia Cason, Corine Astésano, Daniele Schön

To cite this version:

Nia Cason, Corine Astésano, Daniele Schön. Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. Acta Psychologica, Elsevier, 2015, 155, pp.43-50. 10.1016/j.actpsy.2014.12.002 . hal-02446657

HAL Id: hal-02446657
https://hal.archives-ouvertes.fr/hal-02446657

Submitted on 30 Jan 2020

Title: Bridging music and speech rhythm: rhythmic priming and audio-motor training affect speech perception

Author names and affiliations: Nia Cason[a,b*], Corine Astésano[c,d] & Daniele Schön[a,b].

a) Aix-Marseille Université, Institut de Neurosciences des Systèmes, Marseille, France

b) INSERM, U1106, Marseille, France

c) UMR 7309, Laboratoire Parole et Langage, CNRS & Aix-Marseille University, 5 avenue Pasteur, 13006 Aix-en-Provence, corine.astesano@lpl-aix.fr

d) EA 4156, U.R.I. Octogone-Lordat, 5 allées Antonio Machado, 31058 Toulouse Cedex 09, France, corine.astesano@univ-tlse2.fr

*** Corresponding author.** *Email Address:* niacason@yahoo.co.uk

Contact Details:

Nia Cason: niacason@yahoo.co.uk, Tel: 0033 (0) 4 91 38 55 77

Corine Astésano: corine.astesano@univ-tlse2.fr, Tel: 0033 (0) 5 61 50 24 69

Daniele Schön: daniele.schon@univ-amu.fr Tel: 0033 (0) 4 91 38 55 77

Abstract

Following findings that musical rhythm enhances the perception of subsequently presented speech, we investigated whether priming for spoken sentences with metrical rhythms can enhance phonological processing - the building blocks of speech - and to what extent this priming effect is sensitive to rhythmic audio-motor training. Participants heard a metrical prime followed by a sentence (with a matching/mismatching prosodic structure in terms of both stress patterns and the number of elements), for which they performed a phoneme detection task. Behavioural (RT) data were collected from two groups of participants: a group who underwent audio-motor training with the musical rhythms (AM group), and a group who did not (A group).

We hypothesised that 1) inducing metrical expectations about sentence prosody through a musical rhythmic prime (i.e. metrical priming) would enhance the phonological processing of sentences when expectations were matched, and 2) that audio-motor training with musical rhythms would enhance this effect. Indeed, our results show that providing a matching rhythmic prime context resulted in faster phoneme detection, thus revealing a cross-domain effect of musical rhythm on the phonological processing. In addition, our results indicate that rhythmic audio-motor training enhances this priming effect. These results have important implications for rhythm-based speech therapies, and suggest that metrical rhythm in music and speech may rely on shared temporal processing brain resources.

Keywords: Speech, Music, Prosody, Metre, Rhythm

Abbreviations: A group (Auditory modality exposure group); AM group (Audio-motor training group); Reaction Times (RTs).

1. Introduction

Though speech and music may seem distinct from one another, both are hierarchically-organised, rule-based systems whose processing shares a wide range of similarities (Jäncke, 2012; Patel, 2011). Here, the current focus is how cross-domain similarities in the processing of ‘rhythm’ can explain the potential for musical rhythm to impact on speech processing. Rhythm (in both domains) can be broadly defined as the temporal organisation of acoustic events which unfold over time. Two basic properties of rhythm can be said to be its ability to give rise to a sense of ‘beat’ - a series of regular and recurrent psychological events (Cooper & Meyer, 1960) - and ‘meter’, which can be described as an emergent temporal structure that results in a hierarchical organisation of salient and less salient events (London, 2012). The present focus is with metre: how musical metre can inform a listener about speech meter, and the consequences of this in terms of speech processing.

Western music typically has a binary, march-like metre (1**2** 1**2** 1**2**) or a ternary, waltz-like metre (12**3** 12**3** 12**3**). The emergence of metre from a sequence of rhythmic sounds is not only automatic (London, 2012), but its perception also said to be present from infanthood (Hannon & Johnson, 2005). This preference for perceiving metrical patterns has long been acknowledged: even on hearing an isochronous sequence of identical sounds, we automatically project a metrical structure onto them, engendering a perceptual illusion of weak and strong elements (Bolton, 1894).

Similarly in speech, salient and less salient syllables form the metrical patterning of utterances. Though speech does not possess the same degree of regularity as music (Patel, 2008), metrical ‘rules’ allow for a degree of rhythmic predictability in speech; a final stressed syllable marks the end of word groups in French, for example (Hirst & Di Cristo, 1998). We will refer here to stress as to the relative emphasis that is not uniquely signaled by intensity changes, but also importantly by changes in pitch and vowel duration. There are multiple levels of saliencies in speech rhythm: there are different *types* of prosodic stress (e.g. lexical stress, pitch stress, emphatic stress), all of which have different functions in speech, as well as different *degrees* of stress (e.g. primary and secondary lexical stress; high, low or complex pitch stress). As in the music domain, it is the interactions between these multiple levels of stress that allow for the emergence of speech rhythm (Handel, 1989: pp. 383).

In both domains, metre is key in allowing a listener to form predictions about what will happen next. For instance, on hearing a ternary metrical sequence in music (12**3** 12**3** 12**3**), we are able to automatically and implicitly predict that a ‘12**3**’ (weak-weak-strong) pattern will follow. This is also the case in speech, whereby hearing a list of trisyllabic words with a final stress induces expectations for a word with the same metre (Pitt & Samuel, 1990). Given this, it is perhaps not surprising that attentional accounts have been used to explain predictive mechanisms in both speech (Pitt & Samuel, 1990) and in music (Jones, 2008; Large and Palmer, 2002). In speech, The Attentional Bounce hypothesis states that attention is oriented to syllables which are expected to be stressed (Pitt & Samuel, 1990). It claims that the position of these stressed syllables can be predicted on the bases of the metrical patterns of speech, and that this is reflected by quicker phoneme detection at attended syllables (Pitt & Samuel, 1990; Rothermich, Schmidt-Kassow & Kotz, 2011). In music, the Dynamic Attending Theory (DAT) predicts that, on hearing a metrical rhythm (external oscillators), neural rhythms synchronise to it with similar phase and period relations (internal oscillators), and that this coupling dynamically modifies attention in time, with more attentional resources being allocated to strong (predictable) metrical positions (Jones, 2008; Large and Jones, 1999; London, 2012; Snyder and Large, 2005). As a result, auditory events occurring at these more predictable locations are better processed (Barnes & Jones, 2000; Ellis & Jones, 2010; Jones, Boltz & Kidd, 1982; Jones, Moynihan, MacKenzie & Puente, 2002). This effect of auditory rhythmic expectation is also cross-modal, and can enhance the detection of visual targets, too (Brochard et al., 2013; Escoffier et al., 2010).

To conclude thus far, metre can be said to dynamically modulate attention in time in such a way that processing efficiency is influenced. This possibly takes place is a similar way in music and speech, as shown by impacts of musical metre on the phonological processing of speech (Cason & Schön, 2012). Should this be the case, expectations induced by musical metre can be hypothesised to impact on the processing of speech metre, and, as a consequence, to impact on the phonological processing of speech. Notably, this prediction is scale free: it considers only metrical relations, which are considered to include ‘number’ expectations (how many elements to expect) and ‘stress’ expectations (what stress patterning to expect), and not the exact durational expectations (i.e. inter onset intervals).

This hypothesis (that musical rhythm can have a cross-domain impact on speech processing) has been received convincing support. For instance, temporal alignment between stressed beats and stressed syllables allows for a greater comprehension of speech (Gordon, Magne & Large, 2011). More specifically in relation to the current research question, inducing rhythmic expectations can have also have a cross-domain effect on *subsequent* speech processing (Cason & Schön, 2012). In that experiment, expectations about beat and metre were induced by a musical rhythmic prime, which was followed by a bi- or trisyllabic pseudoword with a final stressed syllable. Focussing here on the meter aspect, we found electrophysiological (EEG) evidence of an enhanced processing when the speech metre matched the prime metre (for example, when a prime with a ternary metre (12**3** 12**3** 12**3**) was followed by a trisyllabic word with the same ternary metre (12**3**)). These results indicate that musical rhythm impacts on phonological processing (stimuli for this experiment were pseudowords with no semantic meaning) by inducing domain-general metrical expectations.

To build upon these findings, the current study aimed to investigate whether this effect - seen at the level of single pseudowords - would also be present for real sentences.

Because active (audio-motor) engagement with musical rhythm may allow listeners to further internalise musical rhythms which can be shared by speech, a second aim of the present study was to investigate the role of rhythmic sensorimotor training on this cross-domain priming effect. In both speech and music, audio-motor interactions are vitally important for the consolidation of accurate production and the continual monitoring of output (Guenther, 2006; Lappe, Herholz, Trainor & Pantev, 2008; Levelt et al., 1999; Zatorre, Chen & Penhune, 2007). In music, audio-motor training can strongly influence the perception of rhythm (Phillips-Silver & Trainor, 2005, 2007; Su & Pöppel, 2012) and can enhance listeners' sensitivity to metrical deviants (Geiser, Sandmann, Jancke & Meyer, 2010; Vuust et al., 2005). Considering this, rhythmic training would intuitively result in a stronger priming effect, through an enhanced metrical sensitivity. From a clinical standpoint, the importance of musical audio-motor training in enhancing speech fluency has not been fully acknowledged. In therapies such as Rhythmic Speech Cueing (Thaut, 2005), for example, it is often the therapist who produces the rhythm on a drum whilst the patient is asked only to produce speech (rather than themselves actively participating in the musical rhythm production). In this case, any further benefits of motor rhythmic engagement are not fully potentiated.

In the present experiment, a musical metrical prime sequence was used to induce metrical expectations about both stress patterns and the number of elements. A subsequently-heard sentence either conformed to these stress and number expectations - and thus to listeners' metrical expectations - or did not. Considering the impact of rhythmic predictability on processing efficiency, we hypothesised that in conditions where metrical expectations were met, phonological processing of a target sentence would be enhanced. The experiment employed a phoneme detection task to measure phonological processing, a building block of speech perception, with faster reaction times (RTs) indicating a facilitated access to phonological information. To address the second aim of this study, we tested two groups: one group who underwent audio-motor training with the musical rhythms presented in the experiment (Audio-Motor (AM) Group), and another group without this training (Auditory only (A) Group). We hypothesised that the positive effect of metrical priming on phonological processing would be more striking in the AM Group due to the consolidation of metrical representations afforded by audio-motor training.

2. Materials and Methods

2.1 Stimuli

Experimental trials consisted of a rhythmic prime sequence followed by a sentence. We used four rhythmic prime sequences that comprised four experimental blocks (Prime 1, Prime 2, Prime 3, Prime 4), and four sentence types built around these same rhythms (Figure 1).

Prime sequences were created in Adobe Audition and differed in number of beats (6 or 7) and in the placement of two stressed events. These rhythmic patterns were composed of percussion sounds, which had a stimulus onset asynchrony of 225 msecs (beat level=450ms). The first stressed percussion sound of each prime had a rim shot timbre, a duration of 196 ms and an average Root Mean Square (RMS) of -41.25 dB, the second stressed percussion sound of each prime had a snare timbre, a duration of 353 ms and an RMS of -28.88 dB, and unstressed percussion sounds had a closed high hat timbre, a duration of 138 ms and an RMS of -43.63 dB.

Sentence stimuli were constructed around the four prime rhythms, thus using two syllable groups of 3 or 4 syllables (the most common meter in French, whatever the speaking style, Astésano, 2001). 40 sentences were constructed for each of the four prime rhythms, resulting in a total of 160 sentences. The final word of the sentence (in which the phoneme target was present 50% of the time) was selected using Lexique 3 (New, Pallier, Ferrand & Matos, 2001). These sentence-final words were bisyllabic, had a CV/CV structure (at the phonological level), and were balanced for lexical frequency across the four sentence rhythm types. The last syllable of each word (an open CV syllable) contained a target vowel, for instance, as /i/ in */mari/* for 'mari', and the previous consonant macro-class (liquid, nasal, unvoiced fricative, voiced fricative, unvoiced stop and voiced stop) was balanced over the 4 rhythmic conditions.

Sentences were recorded in a soundproof booth by a native French speaker. Sentences were spoken within a carrier sentence simply to control for sentence-ending prosodic effects; the extracted (first) half of the sentence was followed by a carrier (second) half of the sentence which had exactly the same prosodic structure. To present an example, from *'J'ai bu un bon cafe' n'est pas facile a dire* [*'I drank a good coffee' is not easy to say*], we extracted the first part of the sentence (*J'ai bu un bon cafe*) "x x x X x X"), which has exactly the same prosodic structure as the following (discarded) '*n'est pas facile a dire'* ("x x x X x X"). Without these carrier sentences, the prosodic structure of the sentences may have been confounded, due to phrase-final falling intonation, a well-established phenomenon in French. This characteristic sentence-final stress pattern is marked by lengthening, pitch changes and, to a minor extent, intensity. By recording within carrier sentences, we thus eliminated this final stress. In addition, all sentences were presented in all possible conditions across participants (fully counterbalanced design) thus eliminating any possibility that acoustic differences in syllable saliencies could bias results. Determining syllable onsets and offsets of extracted sentences (PRAAT speech segmentation software, Boersma & Weenink, 2012) allowed duration values to be calculated (sentence duration, final syllable duration, and the onset of the final word relative to the sentence onset). These duration values were used to determine when each sentence would be presented following the prime as well as to timelock RT data to the onset of the final (target) vowel. Examples of the 4 primes and corresponding sentences are available as supplementary audio material.

2.2 Participants

Behavioural (RT) data were collected from two groups of healthy participants. The 'auditory only' group (A Group) was exposed to the rhythmic primes auditorily during the experimental task, whilst the 'audio-motor' group (AM Group) underwent an additional period of audio-motor training with the rhythmic primes. In the A Group, 17 participants (7 female) between the ages of 22 and 39 years (mean = 28 years, 4 months) participated in the experiment. In the AM Group, 17 participants (7 female) between the ages of 23 and 42 (mean = 30 years) participated in the experiment. All participants had normal hearing, were non-musicians and were native French speakers. Behavioural data were collected during a RT (phoneme detection) task. Each participant gave informed consent prior to experimentation and was compensated for their time with a gift.

2.3 Procedure

For both groups (A and AM), participants sat in front of a computer screen. The volume for auditory stimuli was adjusted to a comfortable level. For the duration of each trial, participants were presented with a target vowel on a computer screen and were asked to decide if it was or was not present in the final syllable of the auditorily-presented sentence, and to make their responses as quickly as possible by pressing either a ‘yes’ button or a ‘no’ button. The hand in which each button was held was balanced across subjects.

After a training session of 12 trials, the experiment began. The experiment consisted of 4 blocks of 40 trials each, with each block dedicated to one of the four primes (order counterbalanced across subjects). For each trial, the target vowel (/a/, /e/, /i/, /y/, /u/ or /o/) was present within the heard sentence 50% of the time. These vowel targets were equally distributed across all conditions for each participant. In order to minimise the effects of beat entrainment, the beat relationship between the prime and the sentence was disrupted: the first stressed syllable of sentences landed 2.5 beats (1125 ms) after the final stressed beat of the prime.

As to the manipulation of interest, sentences could either match or mismatch the metrical structure of the prime in terms of both the number of elements and/or the stress pattern (number and stress are two aspects which contribute to ‘metre’); ‘number’ refers to the number of the prime’s percussion sounds compared to the number of syllables in the following sentence, and ‘stress’ refers to the patterning of stressed and unstressed percussion sounds compared to the patterning of stressed and unstressed syllables in the following sentence. Overall, Condition 1 presented a matching metrical structure between the prime and the sentence, and Conditions 2, 3 and 4 presented mismatching metrical structures of varying degrees. More specifically, the experimental conditions were as follows: Condition 1) number and stress match; Condition 2) number match, stress mismatch; Condition 3) number mismatch, *partial* stress match (the first half of the sentence matched the first half of the prime’s stress patterning); Condition 4) number mismatch, stress mismatch (full mismatch). Conditions 2, 3 and 4 therefore each provided different degrees of a metric mismatch between the prime and sentence (see Figure 2 for an example). The order (and frequency) of these 4 experimental conditions was pseudo-randomised within each experimental block, yielding 40 trials per condition for the whole experiment. Participants heard each sentence only once, though sentences were used in all four conditions across participants (i.e. it was a fully balanced design). Error feedback was received at the end of each experimental block. Both visual (target phoneme) and auditory stimuli were presented using Presentation Software.

This protocol was identical for both the A and AM groups. However, the AM group also had a short audio-motor training: they were asked to copy (vocally) the prime rhythm using different sounds to distinguish between strong and weak elements of the rhythmic prime (e.g. ‘ba’ and ‘ka’: a prime of x x X x x X would be repeated as ‘ba ba KA ba ba KA). For each experimental block (i.e. each prime), participants in the AM Group underwent two periods of this audio-motor training, once before the experimental block (n=10 trials), and once midway through (n=10 trials). This made a total of 20 audio-motor rehearsals per prime, which were recorded using a M-Audio Microtrack 24/96 recorder.

For clarity, what follows is a summary of the experimental protocol: there were 4 prime blocks and 4 experimental conditions. For each experimental trial, participants were to decide if a target vowel (presented on-screen) was present in the final syllable of the sentence as quickly as possible, using a ‘yes’/’no’ button press. We expected to see quicker RTs for metrically matching conditions (Condition 1), and that any differences in the mismatching conditions (Conditions 2, 3 and 4) may depend on the type of mismatch (stress/number/partial stress mismatch). In addition, there were 2 groups; one with audio-motor training (AM group) and one without (A Group), the hypothesis being that audio-motor training would augment the cross-domain priming effect, and also that the ability to reproduce the rhythm of a given prime may be related to the subsequent priming effect.

2.4 Data Acquisition and Analyses

Analyses on behavioural RT data were performed for correct trials only. For each participant, outlier responses were removed from data (RTs which were +/- 2.5 x s.d. from the mean). Repeated-measures analysis of variance (RM-ANOVA, mixed factorial design) was used, with the categorical predictor ‘Group’ as a between-subjects factor (2 levels, A Group or AM Group). Within subject factors included ‘Prime’ (4 levels; Prime 1, Prime 2, Prime 3, Prime 4), and ‘Condition’ (4 levels; Condition 1, Condition 2, Condition 3, Condition 4). All p-values reported below were adjusted using the Greenhouse-Geisser correction for nonsphericity, when appropriate, and Fisher least significant difference (LSD) test was used in post-hoc comparisons. Prime reproduction (for the AM group, during training) were scored as follows: 1 for a correct reproduction (the number/stress pattern of the prime sequence was reproduced without errors), 0.5 for a correct reproduction but with a noticeable hesitation, and 0 for incorrect prime reproduction. Of the 17 participants in the AM Group, we collected recordings from 13 participants. The scores for the first 5 repetitions of each prime were taken. Friedman’s test was used for this reproduction analyses.

[Insert Figure 1 around here]

[Insert Figure 2 around here]

3. Results

All participants were able to perform the task easily (mean percentage of incorrect/missed trials in the experimental blocks = 5.88% for A group, 5.55% for AM group; standard deviation across condition was 0.35).

A three-way RM-ANOVA was run on correct response data with 'Group' (A group/AM group) as a between-subjects factor and 'Prime' (Prime 1/Prime 2/Prime 3/Prime 4) and Condition (Condition 1/Condition 2/Condition 3/Condition 4) as within-subject factors.

3.1 The Effect of a Matching Metrical Prime

To consider first the main effects over both groups, there was a main effect of Condition (F(3, 96)=3.5072, p=.0182). A post-hoc analysis revealed this to be the result of quicker RTs in the matching Condition 1 (mean 521.805 msec) compared to those in Conditions 2 (p=0.004, mean 547.08 msec), Condition 3 (p=0.064, mean 538.157) and Condition 4 (p=0.0077, mean 545.605) (metrically mismatching conditions).

Whilst the triple interaction (Group x Prime x Condition) was not significant, we found there to be a strong trend towards significance in the Group x Condition interaction (F(3, 96)=2.3, p=0.08). Post-hoc comparisons showed that Condition 1 (matching) resulted in shorter RTs compared to all three other conditions (p<0.04, p<0.002, p<0.0002, respectively) in the AM group only (Figure 3).

There was also a Prime x Condition interaction (F(9, 288)=2.342, p=.014), revealing that the main effect of Condition was driven by Prime 1 and Prime 2: a post-hoc analysis revealed there to be an experimental effect of Condition only in Prime 1 and Prime 2 (Figure 4). There was no effect of condition for Prime 3 and Prime 4. For Prime 1, Condition 1 (matching) resulted in significantly shorter RTs compared to Condition 2 (p=0.00004), and Condition 4 (p=0.0036), and was not significantly different from Condition 3 (p=0.618). For Prime 2, Condition 1 (matching) resulted in faster RTs compared to condition 2 (p=0.077), Condition 3 (p=0.04) and Condition 4 (p=0.024). Interestingly, results of the behavioural measures during audio-motor rhythm training showed that (for the AM Group), Primes 1 and 2 were also the easiest to reproduce (mean error rates: Prime 1: 1.875 %, s.d. 4.411; Prime 2: 6.736%, s.d. 11.27; Prime 3: 9.556%, s.d. 24.09; Prime 4: 10.417%, s.d. 13.97), although these differences did not reach statistical significance.

[Insert Figure 3 around here]

[Insert Figure 4 around here]

4. Discussion

The aim of this experiment was two-fold. First, we wanted to extend previous findings that musically-induced rhythmic expectations (rhythmic priming) can impact on speech processing of spoken pseudowords (Cason & Schön, 2012). We did this by testing whether rhythmic primes would allow a listener to form rhythmic expectations about a subsequently-presented sentence. Importantly, the rhythmic primes induced metrical expectations - the number and stress patterning of elements - rather than durational (IOI) expectations such as those induced by ‘beat’. Second, we wanted to investigate how additional audio-motor training with the musical rhythms influences this cross-domain effect.

We found that, under certain conditions, sentences preceded by a matching metrical prime were processed more efficiently, as seen by quicker RTs to a phoneme detection task. This effect was more prominent with primes that were more easily reproduced, and when participants were trained to reproduce the prime before the experiment (AM Group). We will first discuss the effect of rhythmic priming on speech processing and will then consider how additional audio-motor training with musical rhythm may enhance the effects of rhythmic priming.

4.1 The effect of metrical priming on sentence processing

In the experimental task, the prosodic structure of heard sentences either matched or mismatched a prior metrical prime in terms of both stress patterning and/or number of elements. No relation was made between durational relations, however (i.e. rhythmic primes were isochronous and sentences were spoken naturally, and so were not isochronous). RTs were taken as a measure of sentence processing efficiency, with quicker RTs to the phoneme detection task taken to indicate a facilitated access to phonological information. Here, we will discuss the finding that metrical predictability of sentences, afforded by musical rhythm, can enhance sentence processing.

Metrical expectations about number and stress were induced. The number of prime sounds was mirrored by the number of syllables in the sentence (Condition 1 and 2) or not (Condition 3 and 4). In the case of stress patterning, the weak and strong prime sounds were fully mirrored by the sentence metre (Condition 1), partially mirrored by the sentence metre (Condition 3) or were not mirrored by the sentence metre (Condition 2 and 4). In partially matching stress patterning (Condition 3), the first half of the sentence conformed to listeners' expectations and diverged from metrical expectations at the 6th syllable, whilst in fully mismatching stress patterning (Conditions 2 and 4), the first stressed syllable occurred either one syllable earlier or one syllable later than anticipated and so diverged earlier, at either the 3rd or 4th syllable. These four conditions were chosen in order to give an indication about which aspects of musical metre (stress pattern, number of elements, or both) might contribute most to inducing metrical expectations. We found that a rhythmic sequence which well-primed for the target sentence (Condition 1) resulted in an enhanced phonological processing of sentences; RTs were significantly quicker than stress mismatching conditions (reaching significance for Conditions 2 and 4, almost significant for the partial stress matching Condition 3). Number and stress are two components thought to comprise metre, and so the finding that partially-matching stress context (Condition 3) resulted in RTs which were more similar to those of the metrically matching condition (Condition 1) could indicate that stress patterning is a more salient metrical cue than number, and that stress expectations which are met in the first half of a sentence may thus also enhance sentence processing somewhat.

The finding that rhythmic predictability can enhance auditory processing is central to the hypotheses presented in this paper. Musical rhythm may induce expectations about 'when' events are expected to occur (Jones & Boltz 1989; Large & Jones, 1999), as well as 'what' events are expected to occur. For instance, on hearing a musical rhythm a listener forms expectations about when in time a future event will occur, and events conforming to these expectations are better perceived in both music (Jones et al., 1982, 2002) as well as speech (Cason & Schön, 2012). The current interest was the ability for rhythmic priming to also induce cross-domain expectations about 'what' speech events will occur (the sequence of weak and strong elements). We reasoned that hearing musical metrical patterns would have the ability to induce expectations about speech metre, and in isolation from durational/beat expectations - the formation of metrical expectations in speech does not require temporally isochronous stimuli (Pitt & Samuel, 1990). Noteworthy, however, is Quené & Port's (2005) findings that whilst temporal regularity in speech (priming for 'beat') enhanced phonological processing, inducing metrical expectations did not. This study differs from our own in several ways, however: firstly, their target stimuli were single words whilst our own were sentences, secondly, we disrupted beat relations between the prime and the sentence so as to measure only the effects of metrical expectancies on sentence processing, and lastly, rather than using speech itself to induce metrical expectations we used a musical rhythmic prime.

While compared to music synchronisation is not afforded to the same degree by natural speech, temporal features of speech are nonetheless vital for perception and comprehension: when speech metre is clearly primed by music, different levels of speech processing may be affected more strongly than speech which is primed by regular speech alone. These different levels may include the phonological level (Cason & Schön, 2012; Pitt & Samuel, 1990; Quene & Port, 2005), syntactic level (Roncaglia-Denissen, Schmidt-Kassow & Kotz, 2013; Schmidt-Kassow & Kotz, 2009), and also semantic level of processing (Rothermich et al., 2011). We would therefore expect the results found here, for phonological processing, to apply to these other levels of speech processing, too. Since no participants claimed to be aware of the rhythmic relationship between the primes and sentences, speech metre may well have been implicitly recognised, much like that of music can be (Schultz et al., 2013); further research would be required to determine whether listeners need to be aware of this metrical relationship for it to impact on their speech processing.

4.2 A differential effect of the four rhythmic primes

There was a difference between primes in their ability to impact on sentence processing: Prime 1 (x x X x x X) and Prime 2 (x x X x x x X) drove the main effect of rhythmic priming on phonological processing; matching conditions (Condition 1) resulted in quicker RTs for these two primes only. We will now discuss why these two rhythms in particular may have been most successful in priming for speech prosody.

Because the four primes were similar in terms of number of sounds, sound durations and tempo, differences should reflect the presence of a more solidified representation of the rhythmic structures. Primes 1 and 2 together differ from Primes 3 and 4 in their initial rhythmic unit: these ‘successful’ primes begin with a ‘x x X’ unit, whilst ‘unsuccessful’ primes begin with a 'x x x X' unit. Bringing the first stress earlier may therefore benefit the induction of metrical expectations. This may be particularly true when using short rhythmical patterns, because one may expect that using longer sequences there would be less metrical ambiguity. There are two non-exclusive possibilities as to why. Considering now Prime 1, it is the only prime with a repeating structure (x x X), whilst the other primes involved no repetition of the same unit. Indeed, the repetition of rhythmic units (in this case, the repetition of 'x x X') is favoured in speech motor rhythms (Pasdeloup, 2005). Listeners' native language (French) may have also played a role in the ability for the rhythmical prime to influence perception; the metric structure of Prime 1 is more representative of speech rhythms typically found in spoken French (Wenk & Wioland, 1982), and thus a more 'familiar' sentence rhythm may have been more prone to disruption in mismatching conditions and/or more prone to facilitating effects in matching conditions.

4.3 The effect of audio-motor training on cross-domain speech processing

Since multimodal engagement with musical rhythm enhances rhythm perception (Geiser et al., 2010; Manning & Schutz, 2013; Su & Pöppel, 2012), it seems intuitive that such training would also impact on speech processing. As noted, while both groups showed a priming effect for the Prime 1 and 2 rhythms, when considering all four prime rhythms, only the AM group showed an effect.

When considering the fact that additional audio-motor training allowed an effect to be seen for all Primes the question arises as to how audio-motor training can increase the priming effect. Matching metre conditions (Condition 1) resulted in quicker RTs to the phoneme detection task than did mismatching metre conditions. Important to note, however, is that the metrically matching condition (Condition 1) did not seem to enhance processing in the AM Group; rather, mismatching metre conditions (Conditions 2, 3 & 4) resulted in slower RTs. That is to say, primes with an incongruent metrical pattern to the target sentence prosody resulted in a more delayed processing for listeners in the AM Group. Nonetheless, this reveals a cross-domain effect of audio-motor training on sentence perception, perhaps through a reinforcement of metrical representations and consequent increase in sensitivity to metrical patterns. This enhanced sensitivity to the prime could result in a greater conflict in mismatching conditions. A similar detrimental processing when attention is allocated to irrelevant information was shown in a study investigating the effect of a background auditory rhythm on visual word recognition (Brochard et al., 2013). Though these are in themselves interesting results, perhaps a greater amount of audio-motor training would be required for a more apparent group difference to be seen. A greater effect of audio-motor training might also be observed if participants are made aware of the rhythmic relationship between the prime and the sentence, as they might be in a therapeutic context.

5. Conclusions

We have replicated previous findings that musical rhythm can induce implicit prosodic expectations, and that this impacts upon the processing of subsequent speech (Cason & Schön, 2012). We have shown this to be the case for real speech (as opposed to pseudowords) and in sentence contexts (as opposed to single words). Moreover, the finding that there is a differential effect of the rhythmical primes used in the experiment suggests that listeners may benefit from shorter, more repetitive rhythmic structures.

Most interestingly, audio-motor training seems to further sensitise listeners to musical rhythm, which in turn increases the potential for musical rhythm to exert an effect on sentence perception. The ability for audio-motor training to amplify the cross-domain effect from music to speech in this way has important implications for using audio-motor musical rhythm engagement in speech rehabilitation contexts, though more research is required to determine how much mismatching music-speech rhythms might interfere with phonological processing, and whether more intense rhythmic training would result in a more potent effect on speech processing.

5.1 Perspectives

Whilst the current experiment has considered the effects of musical rhythmic priming on speech perception, we can also predict that speech production would benefit from rhythmic priming, and through two possible mechanisms. Auditory rhythm engages the motor system (Chen, Penhune & Zatorre, 2008; Grahn & Brett, 2007), and can provide a direct template for timely speech initiation and/or production (e.g. Cummins & Port, 1998). Second, it might be predicted that production may benefit from an enhancement in perception. Several lines of evidence indicate strong links between speech perception and production: speech perception engages areas involved in speech production (review, Galantucci, Fowler & Turvey, 2006; Geiser, Zaehle, Jäncke & Meyer, 2008; Liberman & Mattingly, 1985; Watkins, Strafella & Paus, 2003) which may reflect how the speech production system (motor) is recruited to understand and predict what might happen next (auditory) (Pickering & Garrod, 2007). In this case, an enhancement of perception will necessarily result in an enhanced production, and vice versa. The next step would therefore be to test the effect of rhythmic priming on speech production. This would have more relevance in the context of speech rehabilitation therapies, many of which already recognise musical rhythm as an important fluency-enhancing tool (e.g. Thaut, 2005).

Acknowledgements

This research was supported by a grant from EBRAMUS ITN (Europe BRAin and MUSic) (Grant Agreement number 238157). We thank Frederic Anguille, François Guimont, Bastien Intartaglia, Anaïs Llorens and Patrick Marquis for technical support and help with data collection.

References

Astésano, C. (2001). *Rythme et accentuation en français: invariance et variabilité stylistique*. L'Harmattan.

Barnes, R. & Jones., M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41*, 254–311.

Boersma, P., & Weenink, D. (1999). *Praat: Doing phonetics by computer. Software*. [Computer program]. Version 5.3.05, 〈http://www.praat.org/〉

Bolton; T. L. Rhythm. *American Journal of Psychology,* 1894, 6, 145-238.

Brochard R, Tassin M, Zagar D. (2013) Got rhythm…for better and for worse. Cross-modal effects of auditory rhythm on visual word recognition. *Cognition*, 127(2):214-9.

Cason, N. & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*(12), 2844-2854.

Cooper, G. & Meyer, L. B. (1960). The rhythmic structure of music. Chicago: University of Chicago Press

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, *26*(2), 145-171

Dilley, L.C. & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59,* 294 – 311.

Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Attention, Perception, & Psychophysics*, *72*(8), 2274-2288.

Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. Acta psychologica, 135(1), 12-16.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin & review*, *13*(3), 361-377.

Geiser, E., Zaehle, T., Jäncke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive Neuroscience, 20*(3), 541-552.

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception–training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979-1985.

Gordon, R. L., Magne, C. L., & Large, E. W. (2011). EEG Correlates of Song Prosody: a new look at the relationship between linguistic and musical rhythm. *Frontiers in psychology, 2*.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893-906.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, *39*(5), 350-365.

Handel, S. (1989). Listening: An introduction to the perception of auditory events. pp. 383

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive psychology*, *50*(4), 354-377.

Hirst, D.J. & Di Cristo, A. *(1998).* A survey of intonation systems. *In* D. J. Hirst *&* A. Di Cristo (Eds.), Intonation systems : a survey of twenty languages (pp. 1-44). Cambridge: Cambridge University Press.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *The Journal of the Acoustical Society of America*, *124*, 2263.

Jäncke, L. (2012). The relationship between music and language. *Frontiers in Psychology, 3.*

Jones, M. R. & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Rview, 96*, 459–491.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics, 32,* 211-238.

Jones, M. R., Moynihan, H., MacKenzie, N. & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13,* 313–319.

Jones, M. R. (2008). Musical Time in Oxford Handbook of Music Psychology S Hallam, I Cross, M Thaut Eds, OUP, Oxford.

Kusumoto, K., & Moreton, E. (1997). Native language determines the parsing of nonlinguistic rhythmic stimuli. *The Journal of the Acoustical Society of America, 102*, 3204.

Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *The Journal of Neuroscience, 28*(39), 9632-9639.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychological Review, 106,* 119–159.

Large, E. W. & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26,* 1-37.

Levelt, W. J. (1999). Models of word production. *Trends in cognitive sciences*, *3*(6), 223-232.

London, J. (2012). *Hearing in time*. Oxford University Press.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french: LEXIQUE™. *L'année Psychologique*, *101*(3), 447-462.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36.

London, J. (2012). *Hearing in time*. OUP, USA.

Manning, F., & Schutz, M. (2013). “Moving to the beat” improves timing perception. *Psychonomic Bulletin & Review*, 1-7.

Pasdeloup, V. (2005). Figures et fond dans la scène prosodique: leur résistance face aux variations du débit de parole. In *Symposium Interface Discours-Prosodie-IDP05, Aix-en-Provence, France*.

Patel, A. (2008). *Music, Language and the Brain.* New York: Oxford University Press.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in psychology*, *2*.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430-1430.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533-546.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. *Trends in Cognitive Sciences*, *11*(3), 105-110.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 564–573.

Quené, H. & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica, 62*, 1-13.

Roncaglia-Denissen, M. P., Schmidt-Kassow, M., & Kotz, S. A. (2013). Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *PloS one*, *8*(2), e56000.

Rothermich, Schmidt-Kassow & Kotz (2011). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia, 50,* 232–244.

Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693-1708.

Schultz, B. G., Stevens, C. J., Keller, P. E., & Tillmann, B. (2012). The implicit learning of metrical and nonmetrical temporal patterns. *The Quarterly Journal of Experimental Psychology*, (ahead-of-print), 1-21.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117-126.

Su, Y. H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological research*, *76*(3), 373-382.

Thaut, M. H. (2005). Neurologic Music Therapy in Speech and Language Rehabilitation. *In Thaut M. H. (Ed.), Rhythm, music, and the brain: Scientific foundations and clinical applications (pp. 165–178). New York: Routledge.*

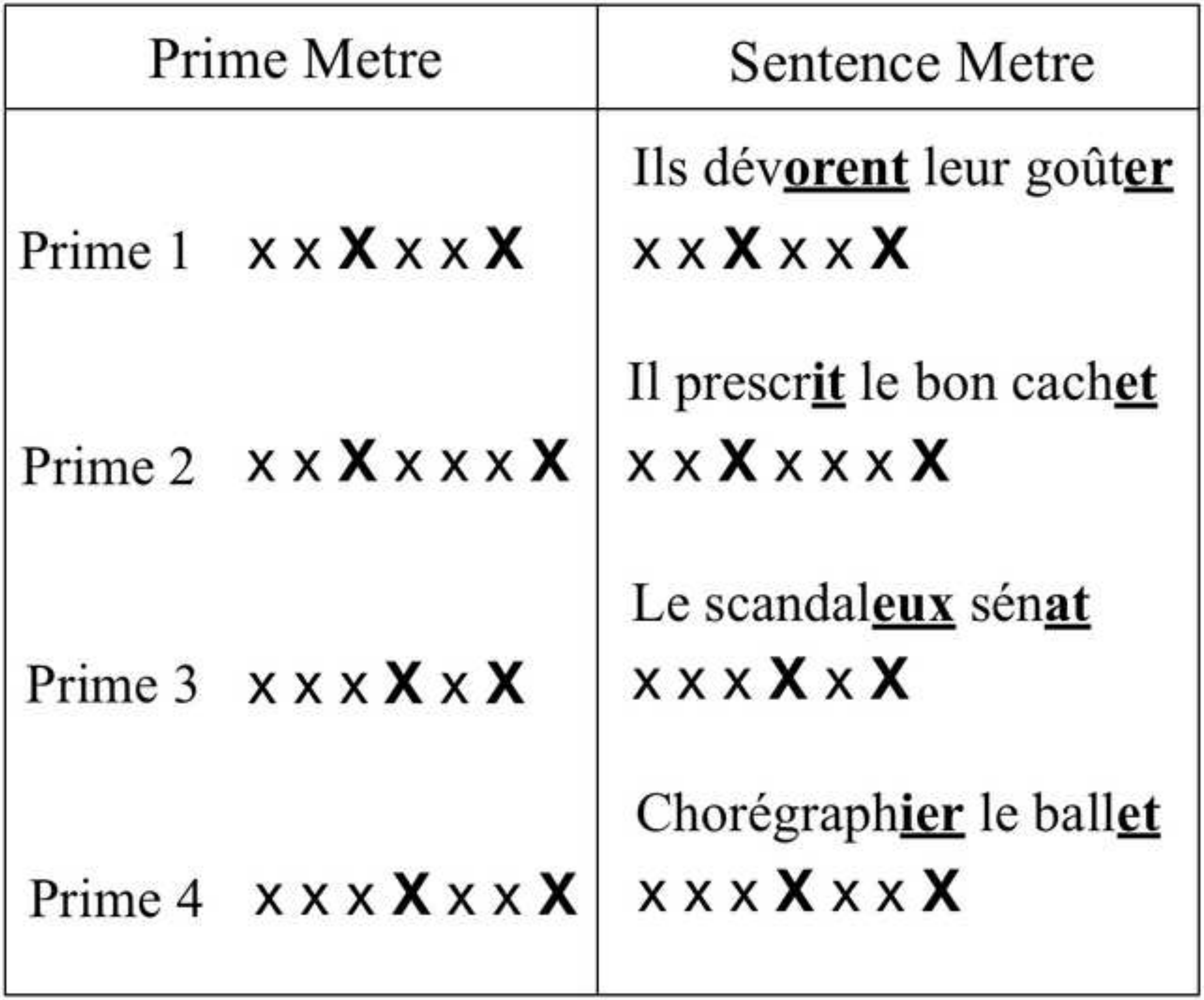
Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter: pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. Neuroimage, 24(2), 560-564.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989-994.

Wenk, B. J., & Wioland, F. (1982). Is French really syllable-timed?. *Journal of phonetics*.

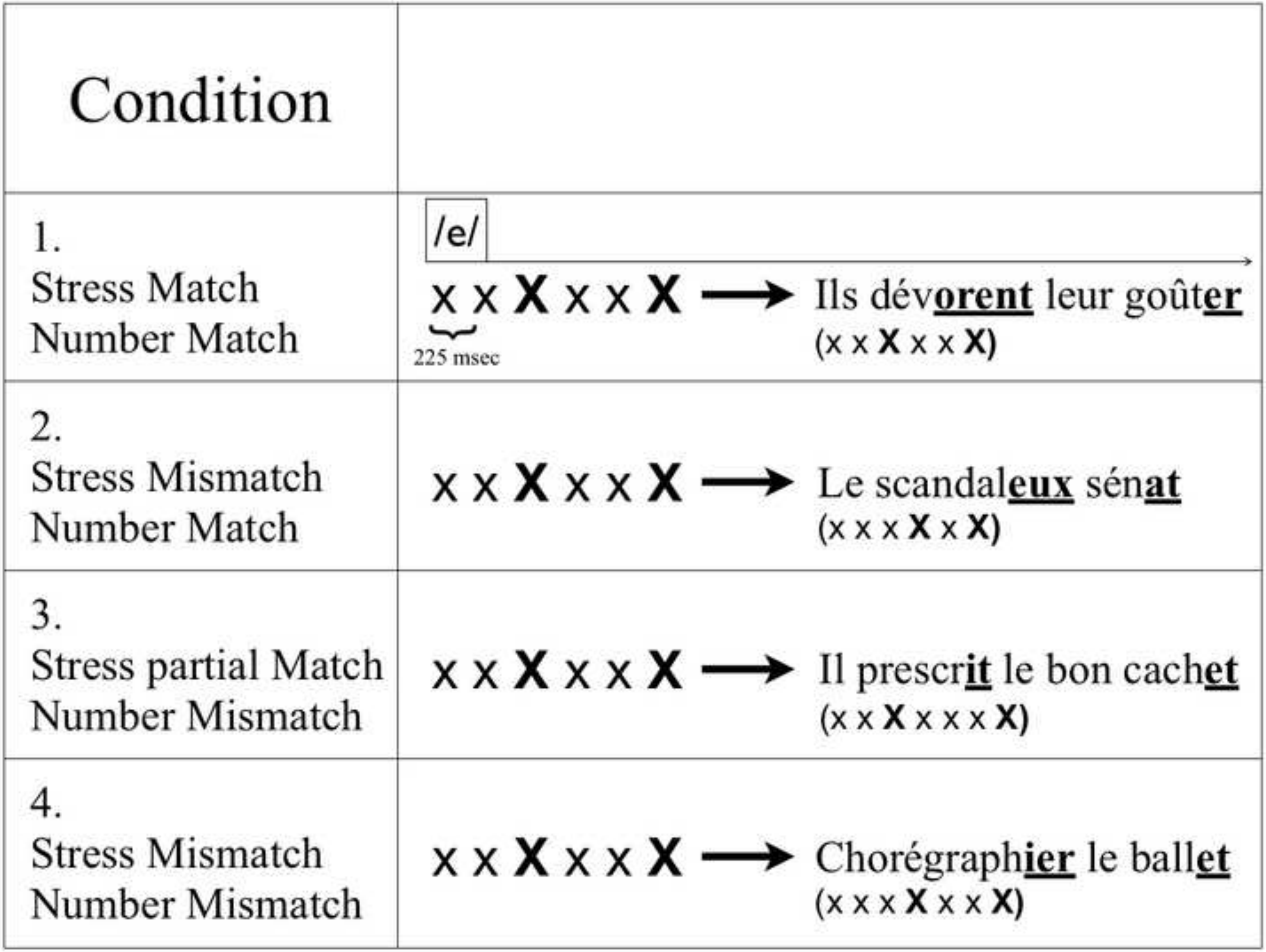
Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547-558.

Figure 1



Prime Metre	Sentence Metre
Prime 1 x x **X** x x **X**	Ils dév**orent** leur goût**er** x x **X** x x **X**
Prime 2 x x **X** x x x **X**	Il prescr**it** le bon cach**et** x x **X** x x x **X**
Prime 3 x x x **X** x **X**	Le scandal**eux** sén**at** x x x **X** x **X**
Prime 4 x x x **X** x x **X**	Chorégraph**ier** le ball**et** x x x **X** x x **X**

Figure 2



Condition	
1. Stress Match Number Match	/e/ x x **X** x x **X** → Ils dév**orent** leur goût**er** (x x **X** x x **X**) 225 msec
2. Stress Mismatch Number Match	x x **X** x x **X** → Le scandal**eux** sén**at** (x x x **X** x **X**)
3. Stress partial Match Number Mismatch	x x **X** x x **X** → Il prescr**it** le bon cach**et** (x x **X** x x x **X**)
4. Stress Mismatch Number Mismatch	x x **X** x x **X** → Chorégraph**ier** le ball**et** (x x x **X** x x **X**)

Figure 3

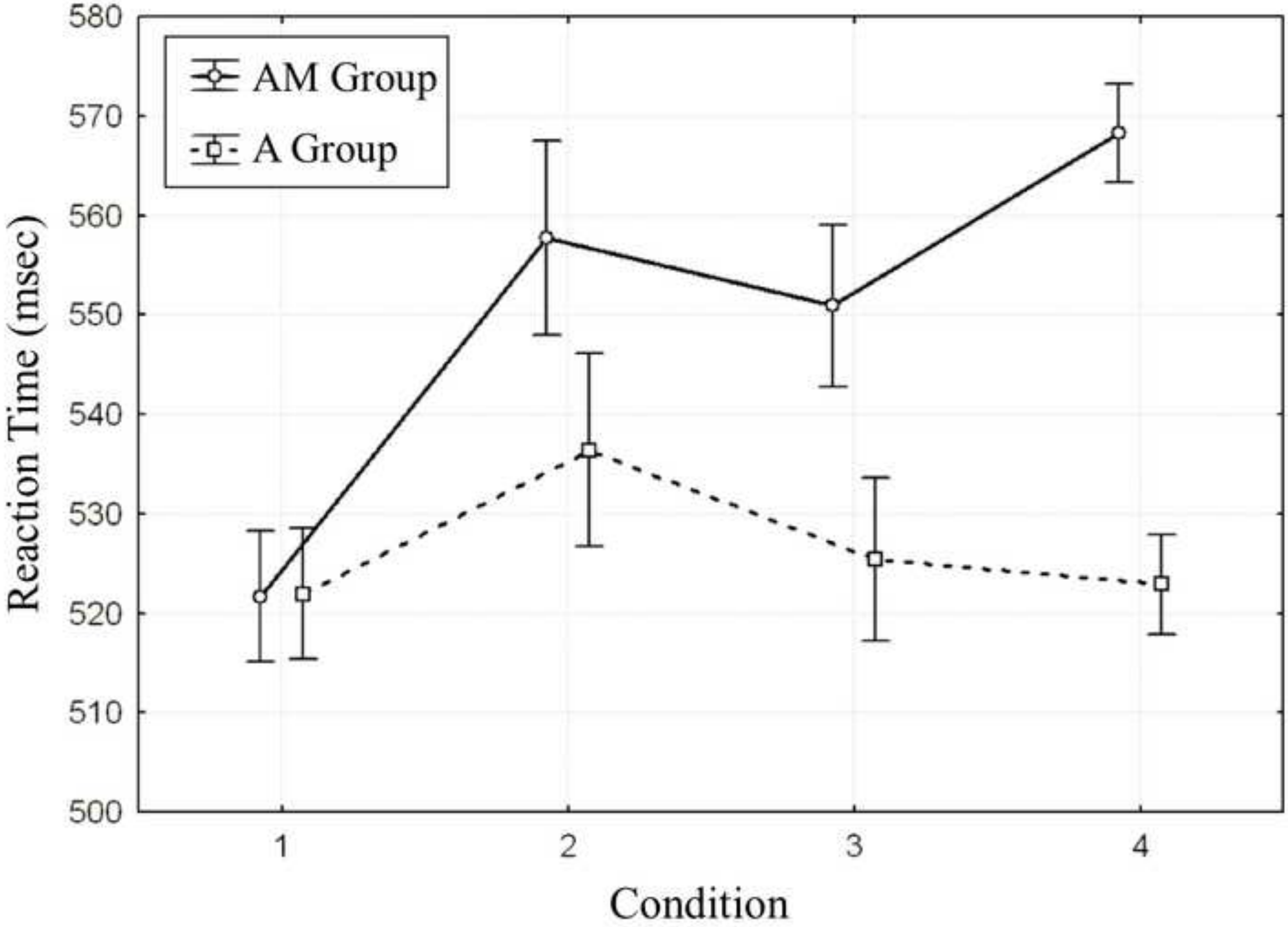

AM Group
A Group
580
570
560
550
540
530
520
510
500
Reaction Time (msec)
1
2
3
4
Condition

Figure 4

Click here to download high resolution image

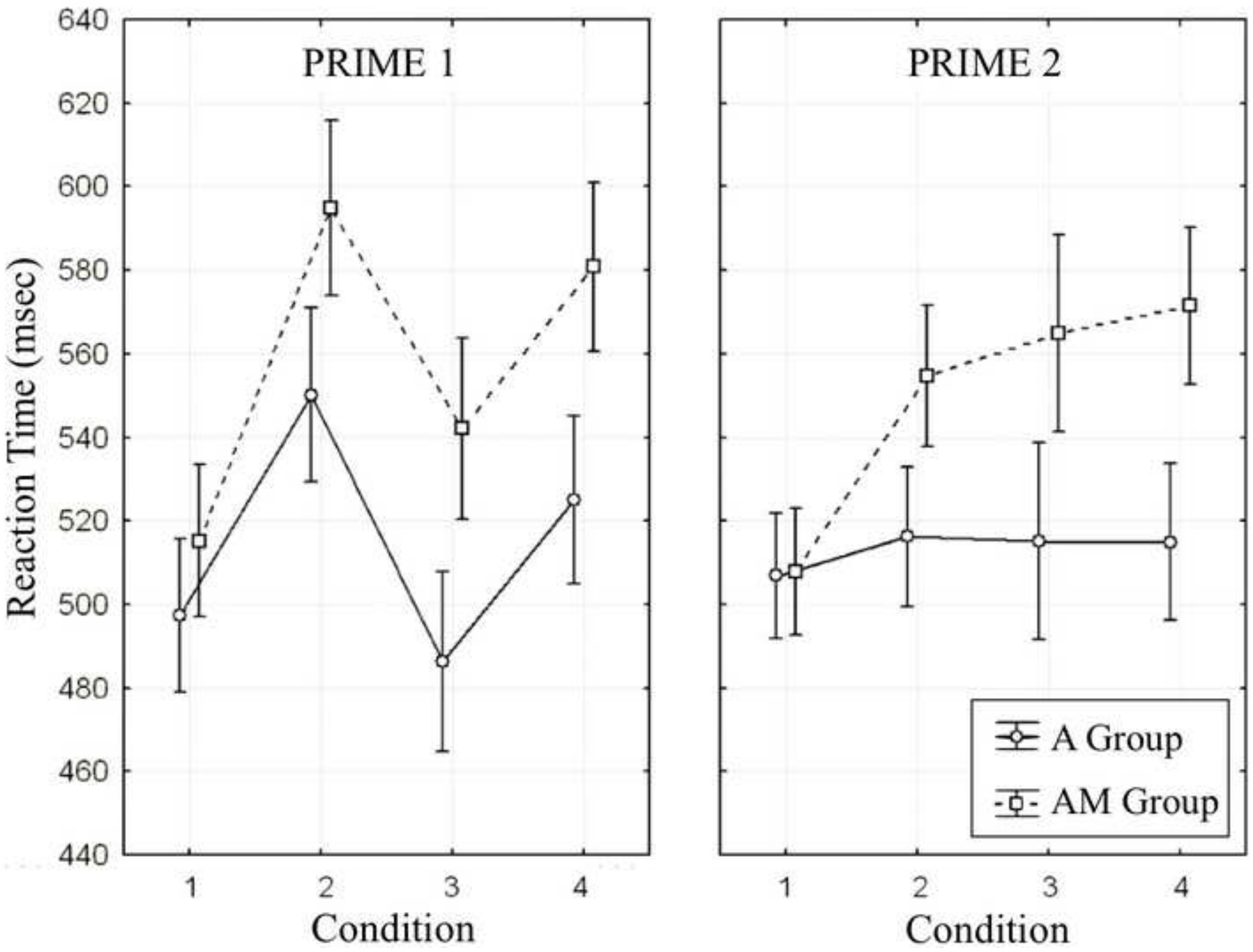


Figure Captions

Figure 1. Schematic representation of the four rhythms used for experimental stimuli. This figure shows each of the 4 prime rhythms and their sentence rhythm ‘equivalents’ (with one example for each). Primes 1 and 2 have an opening of 2 weak beats, Primes 3 and 4 have an opening of three weak beats. Sentences were constructed around these 4 primes and examples are available as supplementary audio material.

Figure 2. The four experimental conditions. Participants heard the prime followed by a sentence, and had to decide whether a target vowel (presented on-screen for the duration of a trial) was present in the final syllable of the word. Prime-sentence pairs were matching/mismatching in terms of both the number and the stress patterning of elements. For instance, in Condition 1, primes were followed by a sentence whose metre was matching both in terms of stress and number. Conditions 2, 3, and 4 were mismatching in terms of stress only (Condition 2), were mismatching in terms of number and partially mismatching in stress (initial stress expectations met) (Condition 3), or were mismatching in terms of both stress and number (Condition 4). The examples presented here are of trials from a Prime 1 block.

Figure 3: Results illustrating RTs for each Condition in the two Groups

Figure 4. ***4a, left:*** Within Prime 1, Condition 1 (matching condition) resulted in faster RTs in both groups. This difference was more pronounced for the AM Group (dotted line). ***4b, right:*** Within Prime 2, Condition 1 resulted in faster RTs in the AM Group only.