Bilingualism: Language and Cognition

cambridge.org/bil

Rhythmic grouping biases in simultaneous bilinguals

Natalie Boll-Avetisyan[1], Anjali Bhatara[2,3], Annika Unger[4], Thierry Nazzi[5,6] and Barbara Höhle[7]

[1]Universität Potsdam, Faculty of Human Sciences, Linguistics Department, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany; [2]Université Paris Descartes, Sorbonne Paris Cité, Paris, France; [3]CNRS (Integrative Neuroscience and Cognition Center, UMR 8002), 45 rue des Saints-Pères, Paris, France; [4]Universität Potsdam, Faculty of Human Sciences, Linguistics Department, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany; [5]Université Paris Descartes, Sorbonne Paris Cité, Paris, France; [6]CNRS (Integrative Neuroscience and Cognition Center, UMR 8002), 45 rue des Saints-Pères, Paris, France and [7]Universität Potsdam, Faculty of Human Sciences, Linguistics Department, Karl-Liebknechtstr. 24–25, 14476 Potsdam, Germany

Research Article

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

Received: 3 July 2018
Revised: 15 January 2020
Accepted: 27 January 2020
First published online: 20 February 2020

Keywords:
simultaneous bilingualism; universal bias; rhythm; rhythmic grouping; Iambic/Trochaic Law

Address for correspondence:
Natalie Boll-Avetisyan,
E-mail: nboll@uni-potsdam.de

Abstract

This study provides a novel approach for testing the universality of perceptual biases by looking at speech processing in simultaneous bilingual adults learning two languages that support the maintenance of this bias to different degrees. Specifically, we investigated the Iambic/Trochaic Law, an assumed universal grouping bias, in simultaneous French–German bilinguals, presenting them with streams of syllables varying in intensity, duration or neither and asking them whether they perceived them as strong-weak or weak-strong groupings. Results showed robust, consistent grouping preferences. A comparison to monolinguals from previous studies revealed that they pattern with German-speaking monolinguals, and differ from French-speaking monolinguals. The distribution of simultaneous bilinguals' individual performance was best explained by a model fitting a unimodal (not bimodal) distribution, failing to support two subgroups of language dominance. Moreover, neither language experience nor language context predicted their performance. These findings suggest a special role for universal biases in simultaneous bilinguals.

1. Introduction

One of the major puzzles in language acquisition research is the following: what cognitive endowment do humans bring to acquire their language(s)? One influential theory is that infants' language acquisition is guided by innate or universal perceptual mechanisms that navigate the processing of speech input to effectively support learning from the start (e.g., Jusczyk & Bertoncini, 1988). In line with this theory, newborns' perception of speech reflects some language-general "universal" categories and biases (e.g., Gervain, Macagno, Cogoi, Peña & Mehler, 2008; Nazzi, Bertoncini & Mehler, 1998; Shi, Werker & Morgan, 1999; Vouloumanos & Werker, 2007). Phonologists assume that phonological structures and patterns that are recurrent across many languages may be a reflection of universal biases on perception, production, and acquisition (for a discussion, see Moreton, 2008).

While universal biases may play an important role at the beginning of first language acquisition, the establishment of the native language system (especially the acquisition of its phonological system) will lead to an adjustment of these language-general categories and biases to language-specific sound properties by maintaining, enhancing or diminishing sensitivity to these properties (e.g., Cutler, Mehler, Norris & Segui, 1983; Kuhl, 2004). In line with this, cross-linguistic studies that have tested potential effects of universal biases on processing have found that their impact on adults' speech perception is modulated by the native language. Effects of this modulation are measurable if the listeners' native language is coherent with a bias; however, if language-specific properties of a native language do not support the maintenance of a universal bias, such effects are not found, or are found to a lesser degree. Examples come from studies on phonotactic universals (e.g., Boll-Avetisyan, 2012 on effects of the Obligatory Contour Principle; Tsuji, Gomez, Medina, Nazzi & Mazuka, 2012 on the effects of the labial-coronal bias; Vroomen, Tuomainen & de Gelder, 1998 on the effects of vowel harmony), and rhythmic universals (e.g., Iversen, Patel & Ogushi, 2008; Bhatara, Boll-Avetisyan, Unger, Nazzi & Höhle, 2013 on the effects of the Iambic/Trochaic Law [ITL] on perceptual grouping). In the present study, we focused on effects of the ITL.

Since language processing procedures are already becoming language-specific during infancy, researchers have limited options at their disposal to test the presence of universal biases in speech perception. They can use neuroimaging techniques with newborns (e.g., Gómez, Berent, Benavides-Varela, Bion, Cattarossi, Nespor & Mehler, 2014), but even

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

newborns' perception may already be influenced by pre-natal experience (Abboub, Nazzi & Gervain, 2016b). They can also measure effects of universal biases in artificial language learning tasks (e.g., Moreton, 2008), but performance in these tasks is also co-affected by linguistic experience (e.g., LaCross, 2015; Mersad & Nazzi, 2011).

The approach the present study uses to test the universality of a perceptual bias has, in the past, received little attention: we look at speech processing in simultaneous bilingual adults who have learned two languages that support the maintenance of this proposed universal bias to a different degree. As we will discuss in more detail below, previous studies have reported strong effects of language dominance or language context in simultaneous bilinguals' speech processing (Cutler, Mehler, Norris & Segui, 1989, 1992; Dupoux, Peperkamp & Sebastian-Gallés, 2010; Sebastián-Gallés, Echeverria & Bosch, 2005). We argue that these modulations within a bilingual group were found because the previous studies tested their perception of speech properties that are not based on a universal bias but that are LANGUAGE-SPECIFIC (i.e., specific vowel or stress categories, and rhythm classes), and, hence, have to be learned exclusively via exposure to the particular language (by means of attending to frequencies and distributions, e.g., Maye, Werker & Gerken, 2002). In this case, frequency of exposure should determine the learning outcomes and, accordingly, language dominance should modulate speech perception. We hypothesize that such within-group modulations by language dominance or language context will not be found for universal biases in bilinguals when at least one of their languages is fully consistent with the bias. We predict, on the contrary, that input from that language would suffice to keep the bias fully active during language acquisition, so that the simultaneous bilinguals will perform at the level of the monolingual group that shows the strongest effects of the bias.

Our test case is a group of French–German simultaneous bilinguals and their rhythmic grouping preferences according to the ITL, as previous studies revealed stronger ITL effects in German- than French-speaking monolingual adults. In the following, we review what is known about ITL processing in monolingual adults and about speech perception in simultaneous bilingual adults. Before that, note that Cutler et al. (1989, 1992) had already proposed that simultaneous bilinguals' processing performance can be informative about universal biases on speech processing. However, their seminal study was ultimately more informative about the acquisition of language-specific processing routines in simultaneous bilinguals than about potential universal processing routines (we discuss their study in more detail below). Since then, the potential for studying effects of universal biases by means of simultaneous bilingual populations has been neglected. We take up this line of research again to investigate whether effects of a universal perceptual bias are resistant or subject to individual variability among simultaneous bilinguals, depending on their language experience, and whether effects of universal perceptual biases are modulated during actual speech processing by language context. Additionally, to shed more light on the role of language experience, we compare the performance of simultaneous bilinguals with that of monolingual speakers of both languages.

1.1. Effects of the Iambic/Trochaic Law

The perceptual bias under study is an assumed innate domain-general auditory principle (Bolton, 1894; Hayes, 1985, 1995; Woodrow, 1909), referred to as the Iambic/Trochaic law (ITL; Bolton, 1894; Hayes, 1995), that has been proposed to explain rhythm perception and to account for an asymmetry in the distribution of rhythm cues in language and music. Typological studies have found large consistencies between languages with regard to the use of rhythmic cues to mark prominent syllables in metrical feet (i.e., the smaller rhythmic units consisting of one or more syllables that make up words): if the foot is trochaic (strong-weak), its prominent initial syllable is typically marked by increased intensity, whereas if the foot is iambic (weak-strong), its prominent final syllable is typically marked by longer duration. Similar asymmetries are found in the distribution of rhythmic cues in music across cultures, where initial beats are marked by higher intensity, and final notes are marked by longer duration in musical phrases (Lerdahl & Jackendoff, 1983; Narmour, 1990; Todd, 1985). Nespor and colleagues (Langus, Seyed-Allaei, Uysal, Pirmoradian, Marino, Asaadi, Eren, Toro, Peña, Bion & Nespor, 2016; Nespor, Shukla, Vijver, Avesani, Schraudolf & Donati, 2008) extended the ITL to account for typological similarities regarding phrasal stress, where trochaic phrasal stress is marked by pitch, and iambic phrasal stress remains marked by lengthening.

The ITL postulates that the similarity of the effects in language and music may be a reflection of a universal bias in auditory rhythmic grouping, according to which listeners would group rhythmically alternating sound sequences into pairs depending on the acoustic cue: listeners perceive iambs (weak-strong) if sounds vary in duration, and trochees (strong-weak) if they vary in loudness or pitch. Many studies carried out over the past century have provided evidence for this law in adults' (Bhatara et al., 2013; Bolton, 1894; Hay & Diehl, 2007; Rice, 1992; Woodrow, 1909, 1951) and young infants' rhythm perception (Abboub, Boll-Avetisyan, Bhatara, Höhle & Nazzi, 2016a, Bion, Benavides-Varela & Nespor, 2011; Hay & Saffran, 2012; Yoshida, Iversen, Patel, Mazuka, Nito, Gervain & Werker, 2010, Molnar, Lallier & Carreiras, 2014), but research with newborns is missing (though see Abboub et al., 2016b, for a first evaluation). There are also studies establishing the ITL in non-human species (birds: Spierings, Hubert & ten Cate, 2017; rats: de la Mora, Nespor & Toro, 2013).

Although the ITL is supposed to be universal, it nevertheless appears to be modulated across languages if language-specific properties do not support the maintenance of the ITL-bias. Studies have revealed cross-linguistic differences suggesting an impact of language-specific experience on this bias in monolinguals (Bhatara et al., 2013; Bhatara, Boll-Avetisyan, Agus, Höhle & Nazzi, 2016; Crowhurst & Teodocio Olivares, 2014; Iversen et al., 2008, Toro, Sebastián-Gallés & Mattys, 2009) that emerges in infancy (Yoshida et al., 2010; Molnar et al., 2014). One factor modulating the ITL cross-linguistically, proposed by Bhatara et al. (2013), is knowledge of lexical stress. They hypothesized that speakers of languages with variable lexical stress such as German draw on abstract phonological representations when perceiving rhythmic speech, while speakers of languages without lexical stress such as French cannot do so, assuming they have no abstract mental representations of stress. This hypothesis was inspired by studies showing that lack of contrastive lexical stress has consequences for stress perception: adult native listeners of languages without contrastive lexical stress show lower accuracy in discriminating pseudowords with different stress patterns than native listeners of languages with contrastive lexical stress (Dupoux, Pallier, Kakehi & Mehler, 2001; Dupoux, Pallier,

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

Sebastián-Gallés & Mehler, 1997; Peperkamp, Vendelin & Dupoux, 2010). Studies with French-learning infants suggest that this relative stress "deafness" emerges early in life, as their ability to discriminate word stress patterns decreases between 6 and 10 months (Abboub, Bijeljac-Babic, Serres & Nazzi, 2015; Bijeljac-Babic, Serres, Höhle & Nazzi, 2012; Höhle, Bijeljac-Babic, Herold, Weissenborn & Nazzi, 2009; Skoruppa, Pons, Bosch, Christophe, Cabrol & Peperkamp, 2013; Skoruppa, Pons, Christophe, Bosch, Dupoux, Sebastián-Gallés & Peperkamp, 2009).

In line with this lexical stress hypothesis, rhythmic grouping was indeed found in both German and French adult listeners, but it was comparatively enhanced in the German listeners (see also Bhatara et al., 2016). Moreover, only German listeners experienced the illusion of hearing trochees in rhythmically invariant speech sequences. This default trochaic grouping procedure in absence of acoustic cues to rhythm must reflect an influence of abstract knowledge of the typical German foot structure.

Another factor modulating the effects of the ITL cross-linguistically, proposed by Nespor and colleagues (2008), is knowledge of phrasal stress (linked to constraints on word order): listeners who have experience with phrase-final heads perceive iambic groupings when hearing sounds with varied duration, as predicted by the ITL; however, if listeners only have experience with phrase-initial heads they perceive trochees even in this condition. This proposal is supported by adult (Iversen et al., 2008; Langus et al., 2016) and infant (Yoshida et al., 2010) studies.

Yet French- and German-learning 7.5-month-old infants did not show differences in rhythmic grouping, which suggests that their grouping corresponds to a language-general bias (Abboub et al., 2016a). Moreover, as a study with French late learners of German showed, learning a second language with variable lexical stress in adulthood can enhance grouping preferences (Boll-Avetisyan, Bhatara, Unger, Nazzi & Höhle, 2016). Together, these studies suggest that effects of the ITL may be present from birth. With the acquisition of the native language, sensitivity to the ITL would decrease when acquiring French, while this initial ITL-bias would be maintained when acquiring German, as the rhythmic structure of German is consistent with this bias. In sum, the ITL seems to be a good candidate for a universal language-general perceptual mechanism that is modulated by language experience. In the present study, we explored its modulation in French–German simultaneous bilinguals. The assumption of the ITL as a universal bias implies that exposure to German from birth would support its maintenance in this group of simultaneous bilinguals independently from the amount of German they receive and thus independently of any effects of language dominance. Before presenting our exact hypotheses, we discuss what previous studies have revealed about speech processing in adult simultaneous bilinguals.

1.2. Speech perception by simultaneous bilinguals

Cutler and colleagues (1989, 1992) were the first to study speech processing in simultaneous bilinguals. Based on previous findings that suggested different language-specific segmentation routines in English and French monolingual speakers (Mehler, Dommergues, Frauenfelder & Segui, 1981; Cutler, Mehler, Norris & Segui, 1983; Cutler & Norris, 1988), they investigated whether both routines are available to simultaneous English–French bilingual speakers, and whether they can apply them selectively depending on the language they are actually exposed to. Their findings at the group level did not resemble either monolingual group. Instead, their analyses suggested that only one language-specific segmentation routine was available to each bilingual, namely the routine that was most efficient for their dominant language (defined as the language the bilingual would like to keep if s/he were to lose one of them). However, contrary to monolinguals, the bilinguals did not apply this segmentation routine when exposed to their language for which this routine was not the optimal one (i.e., the French-dominant bilinguals segmented French but not English based on the syllable), suggesting their ability to abandon the language-specific routine when its application is not appropriate. The authors speculated that in principle (but not in the presented case) the non-dominant language may be processed by reliance on some universal routines, although no specific hypotheses on the nature of such routines were presented.

Following Cutler et al. (1989, 1992), only very few studies investigated speech perception in simultaneous bilingual adults (e.g., Sundara and Polka, 2008), and even fewer explored variability in performance within simultaneous bilingual groups: one was on Spanish–Catalan bilinguals' perception of a Catalan vowel contrast (Sebastián-Gallés et al., 2005) and the other was on French–Spanish bilinguals' stress perception (Dupoux et al., 2010). The perceptual phenomena explored in both cases related to language-specific acquisitions on which non-native monolingual speakers have poorer performance than native monolinguals. Both studies found, as in Cutler et al. (1992), that at the group level, the bilinguals did not pattern with either monolingual group. This was due to individual variability combined with the fact that the bilinguals appeared to be dominant in one language, and that those who were dominant in the language corresponding to the perceptual phenomena explored had higher performance than those who were dominant in the other language. Note that the definition of dominance differed across all three studies: preferred language if having to lose one in Cutler et al. (1989, 1992), the mother's language (Sebastián-Gallés et al., 2005), or some measure of early input (Dupoux et al., 2010). Taken together, these prior studies show that, when it comes to language-specific procedures for which no universal bias is available, and, hence, have to be learned purely from the input, adult simultaneous bilinguals only apply those that are appropriate for their dominant language and thus as a group do not perform at the same level as the corresponding monolingual group.

1.3. Hypotheses and methodological approach

The present study addressed the research question of whether a different pattern of results would be found in simultaneous bilinguals when investigating a universally available routine, the ITL. As discussed above, French-speaking monolinguals show effects of the ITL on their rhythmic grouping, though to a lesser degree than German monolinguals – a result attributable to the fact that the French language supports the maintenance of the ITL to a lesser extent than German (Bhatara et al., 2013). Overall, for simultaneous bilinguals we hypothesized that in such a case their grouping performance should be similar to that of monolinguals of the language for whom more robust, consistent performance is obtained. Hence, we hypothesized that life-long German input from birth should result in robust, consistent ITL-based rhythmic grouping routines, independently of variation in individual

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

exposure to each language or in other factors related to language dominance.

Specifically, we made the following predictions. We first tested whether the ITL is available in bilinguals who have acquired both French and German simultaneously from birth. Given that both French- and German-speaking monolingual populations use intensity and duration for ITL-based rhythmic grouping, we predicted that we should also observe it in simultaneous bilinguals. Moreover, we predicted that simultaneous bilinguals would also perceive trochees when listening to rhythmic invariant speech (like German-speaking monolinguals, but unlike French-speaking monolinguals).

Second, we evaluated how the simultaneous bilinguals compared to the German and French monolinguals tested by Bhatara (2013) using the same procedure. Because effects of the ITL are stronger on German than on French listeners, it is possible that the bilinguals as a group would show intermediate performance between the monolingual groups. Yet, following our hypothesis, we predicted that group performance would be similar to German monolinguals but would differ from French monolinguals.

Third, given that Cutler et al. (1992) found that bilinguals apply only one language-specific procedure (i.e., the one which is appropriate for their dominant language) but restrict its application to the dominant language, we explored whether the degree of ITL-based responses by the bilinguals is modulated by language context, such as the pronunciation of the stimuli as French- or German-sounding, or the language of instruction. If rhythmic grouping by simultaneous bilinguals is modulated by language-context, then stronger ITL-based responses should be found when hearing German-sounding stimuli or when being instructed in German (which may interact with language dominance). However, we predicted that the performance of the bilinguals might not be modulated by these factors, following our hypothesis that life-long German input from birth would generate robust, consistent rhythmic grouping according to the universal ITL.

Fourth, given that prior studies attested clear effects of language dominance on simultaneous bilinguals' speech processing preferences, we tested whether this would also be the case for simultaneous bilinguals' rhythmic grouping preferences (although our prediction was again not to find effects of language dominance). We performed two types of analyses: a) an analysis of the distribution of the individual scores by means of model-based clustering and density estimation, and b) a multiple covariate analysis with potential predictors of language dominance.

Regarding the model-based clustering, Dupoux et al. (2010) demonstrated that the performance of simultaneous Spanish–French bilinguals in prosodic perception was bimodally (rather than unimodally) distributed, with clusters that were highly similar to that of either the Spanish or French monolingual groups. In the present study, we performed similar analyses, but we predicted that the bilinguals' rhythmic grouping performance would be unimodally (rather than bimodally) distributed, since we did not expect the bilinguals' performance to be influenced by language dominance.

Regarding potential predictors of performance, various (and non-consistent) factors (maternal/infancy input, language preference) have been found in past studies to determine language dominance in simultaneous bilinguals. This suggests that bilinguals cannot straightforwardly be subdivided into two categories of dominance, which is in line with theories of bilingualism postulating that differences between bilingualism and monolingualism are gradient (e.g., Baetens Beardsmore, 1982; Grosjean, 1982), depending on a multitude of factors related to language learning and use (e.g., Luk & Bialystok, 2013). For this reason, we opted for a data-driven approach to investigate simultaneous bilinguals' speech processing. That is, rather than defining narrow recruitment criteria *a priori*, we randomly/freely sampled French–German simultaneous bilinguals who had received input in both languages in their first year of life and considered themselves bilingual, and then we used mixed-effects modelling techniques for identifying relevant predictors of bilinguals' performance. Mixed effects models are ideal, as they allow fixed factors to be continuous, and reliably estimate variance if data is not evenly distributed (see, e.g., Baayen, 2010). We did not expect that predictors of language dominance identified in the previous studies would explain French–German simultaneous bilinguals' rhythmic grouping in the same way.

2. Method

2.1. Participants

Thirty-six simultaneous French–German bilingual adults participated. Six of them were excluded, as their age was either below or above the age range (18–40 years) selected in Bhatara et al. (2013). The obtained sample size of 30 participants is justified in two ways: first, based on the effect sizes of prior rhythmic grouping studies that were high (e.g., Bhatara et al., 2013: Cohen's $d = 4.4$ (large) for comparisons between conditions [duration vs intensity, and duration vs. control] within native speakers of German; for comparisons between French and German listeners, Cohen's $d = 1.4$ (large) in the intensity condition, Cohen's $d = 1.1$ (large) in the duration condition). Power analyses for linear mixed effects models are, to date, computationally not possible for experimental designs of the complexity of the current one. However, for an approximation, we performed a post-hoc power analysis of our study using the pangea software (https://jakewestfall.shinyapps.io/pangea/, see Westfall, 2015) that evaluates ANOVA designs. As a second justification for the sample size, this revealed high power for a design including a three-way interaction of condition (3 levels), language of instruction (2 levels) and pronunciation (2 levels) for participant numbers of below 20 (i.e., 5 participants for each combination of language of instruction and pronunciation) with an assumed medium effect size of 0.45 (which is conservative given the large effect sizes found in prior studies) and an alpha error of 0.05. Since we planned to test effects of additional continuous variables, we opted to test a higher number.

Of the 30 participants, nine lived and were tested in Paris, France, and 21 lived and were tested in Berlin or Potsdam, Germany. Inclusion criteria were that they had started receiving input in both French and German (and no additional third language) during their first year of life and could communicate in both languages. They were, however, not recruited for being balanced bilinguals. Instead, we sampled from the full range from French- to German-dominant bilinguals that showed interest in our study.[1] In order to assess their bilingual status,

[1]Contacts with bilinguals were established by means of posts in social media and flyers (distributed at the French embassy in Berlin, in bookstores, and at French–German bilinguals schools), but we also recruited at French–German bilingual events (e.g., *Deutsch-Französischer Stammtisch Berlin*), and from directly addressing people we heard speaking French in Berlin or German in Paris.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

participants filled out a language background questionnaire, based on Dupoux et al.'s (2010) and the LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007).[2] The questionnaire data revealed consistency between participants in that they all had at least minimal exposure to both languages continuously during their lives. They had either grown up in a bilingual family or in a monolingual family with the second language being the majority language spoken in the community. However, they varied regarding the degree of exposure to each language that they had received (for a summary of the questionnaires, see supplementary material Table S1, Supplementary Materials). For this reason, a data-driven approach to analyze effects of experience on bilinguals' performance by means of mixed-effects modelling is ideal. Information from the questionnaire was used to identify potential predictors of rhythmic grouping performance. All participants had normal hearing and no known language disorders. They received a fee for their participation.

2.2. Material

The materials were the same as in the study of Bhatara et al. (2013) and Boll-Avetisyan et al. (2016). Sixteen different CV syllables were constructed by combining four long and tense vowels /e:/, /i:/, /o:/, /u:/ and four consonants of mixed manner and place of articulation /b/, /z/, /m/, /l/. This set of phonemes was selected for two reasons. First, they all are part of both the French and German phoneme inventories. Second, though they may not be perceived the same way by both groups (for example, the German /b/ may sound like a /p/ to the French group; both are voiceless and unaspirated, with short-lag voice onset time) they should nonetheless provide the same variability in segmental material for each group, i.e., the /b/ will not sound the same as the /z/, /l/ or /m/ for any of the group/voice combinations. In each stimulus sequence, each of the 16 syllables was presented twice, once in a strong and once in a weak position. This resulted in 32 syllables per sequence (e.g., /…zu:le:bo:li:lo:zi:mu:be:…/).

Ninety sequences were generated from these syllables. The ordering of the syllables in the sequences was constrained such that they did not contain any syllable reduplications or strings of three identical consonants or three identical vowels. Moreover, we made sure that no CVCV string within a sequence would be a disyllabic word in German or French as listed in the CELEX database (Baayen, Piepenbrock & Gulikers, 1995) or in the Lexique database (New, Pallier, Ferrand & Matos, 2001).

We used text-to-speech synthesis to generate the stimulus sequences because this allows all acoustic parameters to be well-controlled, even across the two languages. For synthesis, we used MBROLA (Dutoit, Pagel, Pierret, Bataille & van der Vreken, 1996), in both a German (De5) and a French pronunciation (Fr4) to control for unintended effects of the language of the voice used. These stimuli resembled artificial language streams used in similar studies (Bion et al., 2011; Saffran, Aslin & Newport, 1996; Tyler & Cutler, 2009). Though the stimuli did not sound like natural speech, they sounded speech-like.

The intensity and duration manipulations were performed by using Praat (Boersma & Weenink, 2010). The F0 contour of all syllables was flat at 200 Hz, a value chosen to be in the middle of the range of F0 for women's spontaneous speech (Baken & Orlikoff, 2000, p. 176). The baseline intensity (mean intensity across the syllable, measured in Praat) was set at 70 dB and the baseline duration at 260 ms for each syllable, 100 ms for the consonant and 160 ms for the vowel. These duration values were chosen based on values reported in previous studies examining stress cues in French and German (Friedrich, Herold & Friederici, 2009; Nazzi, Iakimova, Bertoncini, Frédonie & Alcantara, 2006). The four levels of intensity variation were 2, 4, 6, or 8 dB above baseline and the four levels of duration variation were 50, 100, 150 or 200 ms above baseline. These duration values were larger than those from Hay and Diehl (2007), and the intensity values were smaller. We chose these values based on pilot testing. See Figure 1 for a schematic illustration of the intensity variation applied to the stimuli. All intensity manipulations were applied to the entire syllable, whereas the duration manipulations were applied only to the vowel, given that vowel duration by itself is an important cue in both French and German (Dogil & Williams, 1999; Michelas & D'Imperio, 2010) and one of the main predictors in an automated stress-accent labeling system for English (Greenberg, Carvey, Hitchcock & Chang, 2003). Contrary to Hay and Diehl (2007), there were no pauses between syllables, and all consecutive syllables were co-articulated. To prevent participants from grouping stimuli based on the first pair, the onsets of the stimuli were masked over the first 3 seconds by a combination of white noise, fading out according to a raised-cosine function and fading in of the stimulus, with the intensity increasing also according to a raised-cosine function. As an additional control, half of the sequences began with the strong syllable (longer or louder) and half began with the weak syllable. This control was put in place because Hay and Diehl (2007) reported a strong tendency to group the sequences based on the initial pair of sounds. MATLAB® (R2007b, The MathWorks, Natick, MA) was used to create the white noise and Praat to combine it with the stimuli.

2.3. Procedure

The procedure was identical to that used by Bhatara et al. (2013), the only difference being that the bilinguals were randomly assigned to receiving the instructions in French (15N) or in German (15N). Participants were seated in a quiet room, and the stimuli were presented at a comfortable listening level using PsyScope X (available at http://psy.ck.sissa.it/) on a MacBook laptop. In Potsdam, stimuli were presented through AKG K 55 headphones, and in Paris through Sennheiser HD 558 headphones. Participants were instructed to listen carefully to each sequence and to report whether they heard the alternating stimuli as a strong sound followed by a weak sound or a weak sound followed by a strong sound. They were told that they did not have to wait until the end of a sequence to give their response, but to respond as fast as possible. All of the stimuli were randomly presented within a single block.

Because of the lack of word-level stress in French, instructions given to participants as to what "weak-strong" and "strong-weak"

[2]Dupoux et al. (2010) had also interviewed their participants' parents, however, since in their study, the results from the parents were highly correlated with that of the participants, we only interviewed the participants in our study. In addition, participants completed language proficiency c-tests for both German (Eckes, 2010) and French (received from the Foreign Languages department, University of Potsdam), in which participants filled in missing letters to complete words in texts from both languages in a restricted time window. These test results were, however, not considered further in the analysis, as c-test scores were not found to predict rhythmic grouping performance by L2 learners in our prior study (Boll-Avetisyan et al., 2016). Moreover, the use of c-tests in studies on simultaneous bilingualism is non-standard, because, different from late bilingual acquisition, early bilinguals may have high levels of oral proficiency in both languages which may be uncorrelated with their written proficiency (not having experienced written education), rendering results of these tests irrelevant.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

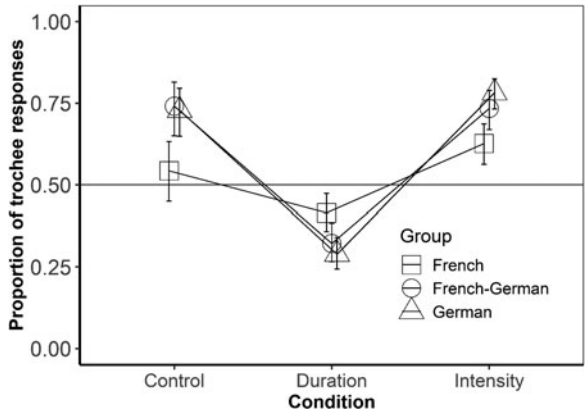


Fig. 1. Model results of the proportion of trochaic responses and their standard errors (back-transformed, but y-axis adjusted to the logit space, 0 = "iambic", 1 = "trochaic") broken down by condition and group.

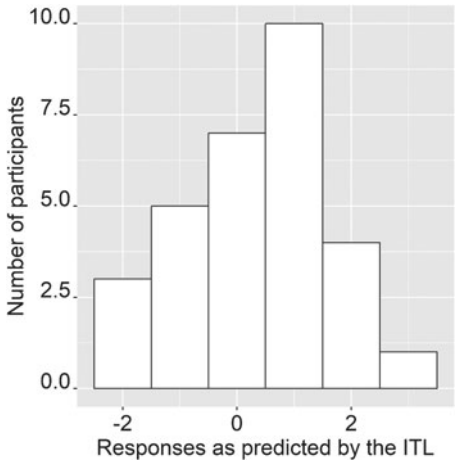


Fig. 2. Histogram illustrating the distribution of the simultaneous bilinguals' individual scores (higher values indicate more ITL-conform rhythmic groupings).

meant differed between groups (receiving French vs. German instruction). The German instruction group was given examples of trochaic and iambic words in German. The French instruction group was given examples of trochaic and iambic words from Spanish as well as examples of contrastive stress in French:

> "You say, 'J'aime le bateau' [I like the boat] and your friend says 'Le gâteau?' [the cake]. You respond, 'Non, le BAteau' placing the emphasis on the first syllable of the word." [This was followed by an equivalent example of emphasis on the second syllable of the word.]

The testing procedure began with four practice trials, two duration-varied sequences and two intensity-varied sequences, both with the maximum variation (8 dB and 200 ms, respectively). Participants pressed one of two labeled buttons to indicate their choice (either a tall bar to the left of a short bar, symbolizing trochaic, or a short bar to the left of a tall bar, symbolizing iambic, see Figure 2), and their responses were recorded.

Over the course of the testing session, participants heard 10 repetitions of each level of intensity or duration variation. Of these 10, five began with a strong syllable and five began with a weak syllable. Participants also heard 10 repetitions of the control sequences. This resulted in a total of 90 stimuli. Left-right position of the response keys was counterbalanced between participants. After they heard 45 stimuli, participants were told they had finished half of the experiment and could take a small break if they wished. Most participants continued immediately with the second half.

Participants were verbally instructed. Of each group – again, randomly assigned – half were exposed to the stimulus set generated with the French pronunciation (n = 16 in the final sample), and the other half to that generated with the German pronunciation (n = 14 in the final sample). After the experiment, participants were interviewed by means of a language background questionnaire that was filled in by the experimenter.

3. Results

Four types of analyses were performed. First, we tested whether the simultaneous bilinguals' grouping preferences differed from chance in the three conditions: Intensity, Duration and Control. Second, the data of the 30 simultaneous bilinguals was compared to the 40 monolingual French and 40 monolingual German listeners tested by Bhatara et al. (2013), to assess whether the rhythmic grouping preferences of the simultaneous bilinguals as a group differed from the other three groups. Third, for exploring the variability among the simultaneous bilinguals, we assessed whether their data was uni- or bimodally distributed. Fourth, we tested whether specific factors related to language experience as well as language context (language of instruction [French vs. German] and stimulus pronunciation [French- vs. German-sounding]) would predict the simultaneous bilinguals' rhythmic grouping.

All analyses were performed in R, version 3.5.3 (R Core Team, 2017). For the mixed-effects models, we used the package "lme4", version 1.1–21 (Bates, Maechler, Bolker & Walker, 2015); and the package "mclust", version 5.4.3 (Scrucca, Fop, Murphy & Raftery, 2016) for the model-based clustering; graphs were generated using the package "ggplot2" (Wickham, 2009).

3.1. Comparisons against chance

We assessed whether grouping preferences by the simultaneous bilinguals were as predicted by the ITL in the three conditions (Intensity, Duration, and Control) by means of a logit mixed-effects model with the intercept set to zero, condition as fixed factor, participants (including a random slope for condition) and items as random factors. The dependent variable response type was binomially distributed (1 = "trochaic" versus 0 = "iambic" responses) and, hence, logit-transformed (i.e., in the model outputs, higher estimate values reflect more trochaic responses). Responses were different from chance in all 3 conditions (all p's < .001; see Table 1 and Fig. 1).

3.2. Simultaneous bilinguals compared to monolinguals

To compare how the simultaneous bilinguals fared in comparison to monolinguals, a global model was calculated that included the present data and data from Bhatara et al. (2013), amounting to 110 participants in total. Fixed factors were condition and group. Participants and items were random factors, and, for participants, a random slope for condition was included. For the two fixed factors, sliding difference contrast coding (an orthogonal contrast available from the MASS package, version 7.3–45, Venables & Ripley, 2002) was used, which assigns the grand

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

Table 1. Parameters of the linear mixed-effects logit regression. The intercept is set to zero, each row indicates the effect of a condition against chance. Level of significance: *p < .05, **p < .005, ***p < .001.

Fixed Effects	*β*	*SE*	*z*	*p*	*Level of significance*
Intensity	1.04	0.18	5.65	< .001	***
Duration	-0.74	0.14	-5.30	< .001	***
Control	1.03	0.22	4.70	< .001	***
Random effects		**Variance**	**Std.Dev.**	**Corr**	
Stimulus (Intercept)		< 0.001	< 0.001		
Participant (Intercept)		0.26	0.51		
Dur–Int		1.78	1.33	-0.63	
Cont–Dur		1.75	1.32	-0.59	-0.90

mean to the intercept. For condition, the contrast was coded to compare duration and intensity (Dur–Int), and duration and control sequences (Cont–Dur) and for group, to compare the simultaneous bilinguals to the French monolinguals (2L1–French) and to the German monolinguals (German–2L1). Model comparisons revealed that an inclusion of the pronunciation factor did not improve the model fit; hence, this factor was excluded.

The model output is provided in Table 2. Estimates (β) indicate the difference scores of the compared levels. Overall (see Fig. 1), participants responded more often "trochaic" when hearing intensity-varied or control sequences than when hearing duration-varied sequences (significant Dur–Int and Cont–Dur, both p's < .001). Comparisons between groups revealed no significant differences between simultaneous bilinguals and German monolinguals, neither in a main effect (2L1–German: $p = .91$) nor in interaction with condition (Dur–Int * 2L1–German: $p = .20$; Cont–Dur * 2L1–German: p = .81).

However, the simultaneous bilinguals significantly differed from the monolingual French: overall, the simultaneous bilinguals gave more trochaic responses than the French monolinguals (significant 2L1–French, $p = .02$). Moreover, the groups differed in both the difference between Intensity and Duration (significant Dur–Int * 2L1–French: $p = .006$) and in the difference between Duration and Control (significant Cont–Dur * 2L1–French): $p < .001$), where the simultaneous bilinguals gave more trochaic responses to both intensity-varied and control sequences when compared to duration-varied sequences than the French monolinguals.

3.3. Analysis of the distribution of individual scores

Next, we explored variability among the simultaneous bilinguals, to determine whether this group would be better described as two groups, reflecting a distribution into French- vs. German-dominant simultaneous bilinguals, or not (as we predicted). For this, we employed the same method as Dupoux et al. (2010): first, we generated a composite score of the participants' responses in each condition (duration, intensity, control) by means of Principal Component Analysis (PCA, see supplementary materials, Table S2 for details, Supplementary Materials).

Next, we applied a model-based clustering method to test whether the distribution of the participants' individual scores (see Fig. 2) would be better captured by a model fitting a unimodal or a bimodal distribution. Model comparisons were based on the Bayesian Information Criterion (BIC; Schwarz, 1978), which rewards if added components significantly account for variance in the data (based on the likelihood function), but, at the same time, penalizes for any added parameter (to lower the risk of overfitting). The results support the assumption of a unimodal distribution of the simultaneous bilinguals' data: the best model was a univariate normal model with one component ($n = 30$, $df = 2$, $BIC = -107.47$). When comparing the groups using the composite score as dependent variable, a linear regression showed significant differences between simultaneous bilinguals and French monolinguals ($\beta = 1.02$, $SE = 0.31$, $t = 3.31$, $p = .001$) but not between German monolinguals and simultaneous bilinguals ($\beta = 0.23$, $SE = 0.31$, $t = 0.76$, $p = .45$), indicating that the distribution of the simultaneous bilinguals' individual performance is indistinguishable from that of the German monolinguals but different from that of the French monolinguals.

3.4. Exploring predictors of simultaneous bilinguals' rhythmic grouping

The analysis included all data of the 30 simultaneous bilinguals for an exploration of the influence of potential predictors of language experience and language context on rhythm perception. Our method was to compare mixed-effects models that either included or excluded predictors to find the combination of predictors that accounted for most variance in the data (as recommended by, e.g., Baayen, 2008; Winter, 2013). Specifically, we considered the model with the lowest Akaike Information Criterion (AIC; Akaike, 1998) value as the best fit of the data. The AIC rewards if added predictors significantly account for variance in the data (based on the likelihood function), but, at the same time, penalizes for any added parameter (to lower the risk of overfitting).

Sliding difference contrast coding was used to assess differences between conditions (i.e., between duration and intensity [Dur–Int], and duration and control sequences [Cont–Dur]), between pronunciations (German–French), and languages of instruction (German–French), the latter two serving as predictors of language mode. As potential predictors of language dominance, we extracted variables from the questionnaire that related to the input participants had received during infant/toddler years, during childhood/teenage years, during recent/current time, and specifically at their homes.[3] Whenever language background questionnaire data from different questions tapped into the same type of information and were highly correlated, we calculated a composite score of these variables by means of Principal Component Analysis (PCA). This procedure was chosen to reduce collinearity in the model, and to reduce the number of tested predictors. A Principal Component that bundles information from correlated factors that relate to one theme is likely to be a stronger representative of a predictor than had we selected one single variable as a representative (see supplementary materials, Table S3 for

[3]Following up on Cutler et al. (1989), we were interested in assessing the potential role of simultaneous bilinguals' emotional bonds with their languages. However, two questions of our questionnaire, which we thought to use for this, were uninformative because of missing values: our participants had large difficulties answering Cutler et al.'s (1992) question about which L1 they preferred keeping if they had to lose one after a stroke. Four of 30 participants refused to tell, and remaining responses were skewed, as 20 participants said they would keep French. We also asked about the language they use with their pets, but 8 participants never had pets.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

Table 2. Parameters of the linear mixed-effects logit regression modelling the comparisons between groups. Level of significance: *p < .05, ** p < .005, ***p < .001.

Fixed Effects	β	SE	z	p	Sign.
Intercept	0.34	0.06	5.94	< .001	***
Dur–Int	-1.60	0.13	-12.09	< .001	***
Cont–Dur	1.40	0.15	9.41	< .001	***
2L1–French	0.32	0.14	2.31	< .05	*
German–2L1	0.02	0.14	0.11	.91	
Dur–Int * 2L1–French	-0.90	0.33	-2.71	< .01	**
Cont–Dur * 2L1–French	1.28	0.37	3.46	< .001	***
Dur–Int * German–2L1	-0.42	0.33	-1.27	.20	
Cont–Dur * German–2L1	0.09	0.37	0.25	.81	
Random effects		**Variance**	**Std.Dev.**	**Corr**	
Stimulus (Intercept)		0.03	0.18		
Participant (Intercept)		0.24	0.49		
Dur–Int		1.58	1.26	-0.48	
Cont–Dur		1.60	1.27	-0.67	-0.92

Table 3. Parameters of the linear mixed-effects logit regression modelling the simultaneous bilinguals' rhythmic grouping preferences.

Fixed Effects	β	SE	z	p	Level of significance
Intercept (Grand mean)	0.44	0.10	4.21	< .001	**
Dur–Int	-1.78	0.26	-6.76	< .001	***
Cont–Dur	1.75	0.29	6.08	< .001	***
Current exposure	-0.09	0.05	-1.96	.05	
Random effects		**Variance**	**Std.Dev.**	**Corr**	
Stimulus (Intercept)		< 0.01	< 0.01		
Participant (Intercept)		0.22	0.46		
Dur–Int		1.79	1.34	-0.68	
Cont–Dur		1.72	1.31	-0.50	-0.90

details, Supplementary Materials). All included fixed factors were coded such that the intercept reflects the grand mean: continuous fixed factors were centered around their mean, and for the categorical fixed factors, sliding difference contrast coding assigns the grand mean to the intercept by default.

By means of this data processing method, eight predictors were yielded that we included in the model comparisons. Regarding input during infant/toddler years, there were high correlations in the self-estimated amount of German input (in percent) received between the ages of 0 to 1 year, 1 to 2 years and 2 to 4 years; hence, one predictor was a composite variable of these data. An uncorrelated variable was the language spoken by the mother to the child, which was used as a second predictor. Moreover, the data regarding the place of residence was correlated at 3 ages (between 0–1, 1–2, and 2–4 years), and was thus merged through PCA to constitute a third predictor.

Regarding input during childhood/teenage years, there were high correlations in the self-estimated amount of German input received at the ages of 4 to 10 years and 10 to 18 years; hence, they were combined through PCA to yield a fourth predictor. Regarding input during recent/current time, we ran a PCA yielding a fifth predictor combining correlated measures regarding current country of residence and self-estimated amount of exposure to German during the past five weeks, the past five months and the past five years. A sixth predictor was used to assess the role of the family language: this three-level factor indicated whether participants' families used solely German, solely French or both languages among each other for communication at home.[4] Together with the two factors for language context (pronunciation and language of instruction), this resulted in eight potential predictors.

The influence of the eight predictors was assessed as follows: we started with a minimal model, which just included condition in the fixed part. Next, we tested models adding individual predictors, and ultimately tested 2- and 3-way interactions of individual predictors and condition. The best model obtained was minimal: the only predictor that improved the model fit was the predictor regarding German language input during recent/current time (henceforth: current exposure) as a main effect ($AIC = 3182.1$). All other predictors did not improve the model fit, and were, hence, not included in the final model.

Results are displayed in Table 3. The main effect of current exposure approached significance ($p = .05$): the less participants were currently exposed to German, the more trochaic responses they gave in the experiment.

4. Discussion

This study investigated the effects of an assumed universal, language-general perceptual bias in simultaneous bilinguals. For this, we explored ITL-based processing by French–German simultaneous bilinguals, predicting that its expression, both in terms of overall group level performance and modulation within the group, should show a different pattern than has been found for bilinguals' acquisition of language-specific processing procedures not based on universal biases (Dupoux et al., 2010; Sebastián-Gallés et al., 2005; Cutler et al., 1989, 1992). The results of our study support our hypothesis that universal biases influence adult simultaneous bilinguals' speech processing irrespective of language experience and dominance.

First, tests against chance revealed that French–German simultaneous bilinguals show consistent iambic grouping preferences when listening to speech streams, in which the length of syllables alternates. When listening to speech with syllables alternating in intensity and when listening to rhythmically invariant speech, they showed trochaic grouping preferences. These results establish that ITL effects are present in simultaneous French–German

[4]Sebastián-Gallés et al. (2005) used a similar factor but specified slightly differently as a score based on sibling-participant- and father-mother-interaction without including mother-participant- or father-participant-interaction. In their study, mother-participant-interaction was a balanced factor.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

bilinguals. Furthermore, they suggest that simultaneous bilinguals apply the same default trochaic grouping procedure in the absence of acoustic cues to rhythm that had previously been found for monolingual German- but not for monolingual French-speaking listeners (Bhatara et al., 2013).

Second, as predicted, the simultaneous bilinguals' performance at the group level was indistinguishable from that of the German-speaking monolinguals but differed from the French-speaking monolinguals. Specifically, simultaneous bilinguals showed stronger, more consistent ITL-based grouping preferences than the French monolinguals in the contrast between the Intensity and Duration condition as well as between the Duration and Control condition (section 3.2.). The overall lack of differences between the simultaneous bilinguals and the German monolinguals is a result that differs from prior studies in which simultaneous bilinguals performed like monolinguals of one or the other language, resulting in intermediate performance at the group level (Cutler et al., 1989; Dupoux et al., 2010; Sebastián-Gallés et al., 2005). This supports our hypothesis that a universal bias is fully expressed in adult simultaneous bilinguals.

Third, given Cutler et al.'s (1989, 1992) finding that bilinguals' use of language-specific segmentation procedures was sensitive to language context, we explored whether this would also be true for processing based on a perceptual bias that is assumed to be universal but is modulated cross-linguistically. As we used artificial nonsense speech streams as material, we investigated effects of language context by testing whether simultaneous French–German bilinguals grouping procedures are modified by language of instruction or language pronunciation (using different voices of the mbrola speech synthesizer). Some studies on non-simultaneous bilinguals' speech perception had reported effects of language of instruction (e.g., de la Cruz-Pavía, Elordieta, Sebastián-Gallés & Laka, 2015; Elman, Diehl & Buchwald, 1977; Soares & Grosjean, 1984), but for simultaneous bilinguals, such an effect is, to this point, unattested. Moreover, non-simultaneous French–English bilinguals' speech perception has previously been found to be affected by whether the pronunciation of the non-word stimuli was French- or English-sounding (Gonzales & Lotto, 2013). Our study, which is the first to test both the influence of pronunciation and language of instruction in simultaneous bilinguals, suggests that there was none. There was neither a general effect of language context, nor an effect of language context modulated by language dominance. We suggest that these factors had no effect because participants relied on a universal processing routine. However, since the language context modulation in the present study differed from that by Cutler et al. (1989, 1992), who presented real English and French words as stimuli, future studies will have to further consider effects of language context to get a better understanding of how simultaneous bilinguals adapt their processing routines to the perceived speech.

Fourth, we found that ITL-based grouping is rather immune to language dominance and is not modulated by specific conditions of simultaneous bilingual language experience. In our study, individual variability in grouping performance among the simultaneous bilinguals was best described by a unimodal distribution, suggesting that the overall performance does not reflect two subgroups of French- versus German-dominant participants (section 3.3). This differs from Dupoux et al. (2010), who demonstrated that lexical stress perception performance by their French–Spanish simultaneous bilinguals was clearly bimodally distributed, suggesting two groups that were dominant in either one or the other language. Moreover, while an effect of early language experience had been previously found (Dupoux et al., 2010; Sebastián-Gallés et al., 2005), our analyses did not provide evidence that simultaneous bilinguals' grouping preferences were modulated by early input, early country of residence, or the language used by the mother. This further supports our proposal that the ITL is a universal processing routine, and that effects of universal perceptual biases are more resistant to variations in the conditions of early bilingual experiences than language-specific routines that are acquired during language development. The only indication of an effect of experience was related to current exposure: bilinguals with more current exposure to German were less likely to perceive trochaic groupings. However, this unpredicted effect only approached significance, so that we are hesitant to consider it in more detail. Hence, language experience is not robustly accounting for individual variability among simultaneous bilinguals in our study. The source of variability that exists in their grouping preferences (see Fig. 2) remains unexplained at this point, but may be due to other cognitive factors such as, for example, auditory acuity (see Boll-Avetisyan, Bhatara & Höhle, 2017, who identified that individual variability in musical rhythm acuity predicted rhythmic grouping of speech), which will have to be investigated in future research.

To sum up, we take the above pattern of findings as new, original evidence that the ITL is a universal bias, rather than the expression of language-specific procedures that are developed during language acquisition. Indeed, one could have argued, for example, that our simultaneous French–German bilinguals use the rhythm-processing routine that they have available for German to process all stimuli, German and French. However, if this were the case, the same should be expected for simultaneous French–Spanish bilinguals' discrimination of stress patterns that are relevant in Spanish but not in French. However, in that instance, clear effects of language dominance on simultaneous bilinguals' processing were found (Dupoux et al., 2010). We argue that, in this prior study, perception relied on language-specific knowledge, namely that of the dominant language in the case of simultaneous bilingualism. For rhythmic grouping, however, a universal bias is available. Hence, French–German simultaneous bilinguals can rely on this bias, irrespective of language dominance as their early exposure to German had supported the maintenance of the bias while the bias was diminished in the monolingual listeners of French as an effect of having acquired a language that does not support its maintenance. Of course, additional data will be needed to further test our interpretation of the ITL as a universal bias, such as more data on the expression of the ITL in very young infants, following up on Abboub et al. (2016b)'s newborn study, and using concatenated speech sequences as those used here rather than syllable pairs.

A question that remains to be explored in the future is the exact role of the rhythmic/prosodic structure of the two native languages on the strength of the effects of the ITL on their general rhythmic grouping. Cutler et al. (1992) proposed that a universal processing routine will always surface in simultaneous bilinguals' processing of their non-dominant language. In the present study, we focused on two languages for which it had previously been established that monolingual speakers of both languages show effects of the ITL on their rhythm perception, although rhythmic grouping preferences were weaker in French than in German monolinguals. What we do not know is how simultaneous bilinguals would perform when either both of their languages weakly support the maintenance of the ITL, or if one of them goes

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

against it (as for the use of duration-based variation in speakers of Japanese, Turkish, Persian or Basque, for which more trochaic groupings of duration-varied sequences was found other than predicted by the ITL; e.g., Iversen et al., 2008; Langus et al., 2016; Molnar et al., 2014). If one language shows properties that contradict the ITL, two scenarios are possible. The first possibility is that simultaneous bilinguals will, in this case, also perform as the monolingual peers of their language supporting a maintenance of the ITL. This result would provide an even stronger case that language dominance and language context may have little effect on performance if universal biases are available. The second possibility is that, in this case, language dominance and context will affect performance. A reliance on the ITL would literally IMPEDE the processing of the language going against the ITL by leading to false rhythmic groupings. Hence, bilingual speakers with dominance in the language going against the ITL may acquire language-specific rhythm processing procedures and suppress the ITL in the context of their dominant language. This issue should be addressed in future studies.

To conclude, we take the present study as showing that simultaneous bilinguals' speech processing can be a test case for studying effects of universal biases on speech processing. In the present case, we provided novel evidence to support the assumption of the ITL as a language-general, universal bias on rhythm perception. The present results establish that universally-guided speech perception routines are resistant to specific input conditions in simultaneous bilingualism and not modulated in the same way as found for language-specific processing procedures in previous studies (Dupoux et al., 2010; Sebastián-Gallés et al., 2005; Cutler et al., 1989). These findings call for future work on simultaneous bilinguals' acquisition of a range of further perceptual processes.

5. Supplementary Materials

For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728920000140.
Supplementary Materials comprise:
Table S1: Participants' background information
Table S2: Results of a Principal Component Analysis over participants' rhythmic grouping data
Table S3a-S3d: Results of Principal Component Analyses over participants' experience factors

Acknowledgements. We thank Wiebke Bruchmüller, Carina Hoppe, Silke Schunack, and Alexandra Schmitterer for help with recruiting and testing participants. This work was supported by two Agence Nationale de la Recherche - Deutsche Forschungsgemeinschaft grants awarded to BH and TN [grant numbers #09-FASHS-018, HO-1960/14–1] and to Ranka Bijeljac-Babic and BH [grant numbers HO-1960/15–1, ANR-13-FRAL-0010].

References

Abboub, N, Boll-Avetisyan, N, Bhatara, A, Höhle, B and Nazzi, T (2016a). An exploration of rhythmic grouping of speech sequences by French- and German-learning infants. *Frontiers in Neurosciences*, **10**:292.

Abboub, N, Nazzi, T and Gervain, J (2016b). Prosodic grouping at birth. *Brain and Language*, **162**, 46–59.

Abboub, N, Bijeljac-Babic, R, Serres, J and Nazzi, T (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, **132**, 111–120.

Akaike, H (1998). Information theory and an extension of the maximum likelihood principle *Selected Papers of Hirotugu Akaike*. New York: Springer, pp. 199–213.

Baayen, RH, Piepenbrock, R and Gulikers, L (1995). *The CELEX Lexical Database* (CD-ROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Baayen, RH (2008). *Analyzing Linguistics Data: a practical introduction to statistics*. Cambridge: CUP.

Baayen, RH (2010). A real experiment is a factorial experiment? *The Mental Lexicon*, **5**, 149–157.

Baken, RJ and Orlikoff, RF (2000). *Clinical measurement of speech and voice*, p. 610. San Diego, CA: Singular Pub Group.

Baetens Beardsmore, H (1982). *Bilingualism: Basic Principles*. Clevedon: Tieto.

Bates, D, Maechler, M, Bolker, B and Walker, S (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bhatara, A, Boll-Avetisyan, N, Agus, T, Höhle, B and Nazzi, T (2016). Language experience affects grouping of musical instrument sounds. *Cognitive Science*, **40**, 1816–1830.

Bhatara, A, Boll-Avetisyan, N, Unger, A, Nazzi, T and Höhle, B (2013). Native language affects rhythmic grouping of speech. *Journal of the Acoustical Society of America*, **134**, 3828–3843.

Bijeljac-Babic, R, Serres, J, Höhle, B and Nazzi, T (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLoS One*, **7**, e30843.

Bion, RAH, Benavides-Varela, S and Nespor, M (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, **54**, 123–140.

Boersma, P and Weenink, D (2010). *Praat: doing phonetics by computer* [Computer program]. Version 5.2.05, retrieved 10 October 2010 from http://www.praat.org/

Boll-Avetisyan, NAT (2012). *Phonotactics and its acquisition, representation, and use: an experimental-phonological study*. Doctoral dissertation. Utrecht: LOT Publications. LOT Dissertation series 298.

Boll-Avetisyan, N, Bhatara, A, Unger, A, Nazzi, T and Höhle, B (2016). Effects of experience with L2 and music on rhythmic grouping by French listeners. *Bilingualism: Language and Cognition*, **19**, 971–986.

Boll-Avetisyan, N, Bhatara, A and Höhle, B (2017). Effects of musicality on the perception of rhythmic structure in speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, **8**, 1–9.

Bolton, TL (1894). Rhythm. *American Journal of Psychology*, **6**, 145–238.

Crowhurst, MJ and Teodocio Olivares, A (2014). Beyond the Iambic-Trochaic Law: the joint influence of duration and intensity on the perception of rhythmic speech. *Phonology*, **31**, 51–94.

Cutler, A, Mehler, J, Norris, D.G and Segui, J (1983). A language-specific comprehension strategy. *Nature*, **304**, 159–160.

Cutler, A and Norris, DG (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 113–121.

Cutler, A, Mehler, J, Norris, D and Segui, J (1989). Limits on bilingualism. *Nature*, **340**, 229–230.

Cutler, A, Mehler, J, Norris, D and Segui, J (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, **24**, 381–410.

de la Cruz-Pavía, I, Elordieta, G, Sebastián-Gallés, N and Laka, I (2015). On the role of frequency-based cues in the segmentation strategies of adult OV-VO bilinguals. *International Journal of Bilingual Education and Bilingualism*, **18**, 225–241.

De la Mora, DM, Nespor, M and Toro, JM (2013). Do humans and non-human animals share the grouping principles of the Iambic-Trochaic Law? *Attention, Perception, and Psychophysics*, **75**, 92–100.

Dogil, G and Williams, B (1999). The phonetic manifestation of word stress. In van der Hulst H (ed), *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 273–334.

Dupoux, E, Pallier, C, Kakehi, K and Mehler, J (2001). New evidence for prelexical phonological processing in word-recognition. *Language and Cognitive Processes*, **16**, 491–505.

Dupoux, E, Pallier, C, Sebastián-Gallés, N and Mehler, J (1997). A destressing 'deafness' in French? *Journal of Memory and Language*, **36**, 406–421.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

Dupoux, E, Peperkamp, S and Sebastian-Gallés, N (2010). Limits on bilingualism revisited: stress 'deafness' in simultaneous French–Spanish bilinguals. *Cognition*, **114**, 266–275.

Dutoit, T, Pagel, V, Pierret, N, Bataille, F and van der Vreken, O (1996). *The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes.* Paper presented at the Proceedings of The Fourth International Conference on Spoken Language Processing (ICSLP'96) 3.

Eckes, T (2010). Der Online-Einstufungstest Deutsch als Fremdsprache (onDaF): Theoretische Grundlagen, Konstruktion und Validierung. In Grotjahn R (ed), *Der C-Test: Beiträge aus der aktuellen Forschung/The C-test: Contributions from current research*. Frankfurt am Main: Lang, pp. 125–192.

Elman, JL, Diehl, RL and Buchwald, SE (1977). Perceptual switching in bilinguals. *The Journal of the Acoustical Society of America*, **62**, 971–974.

Friedrich, M, Herold, B and Friederici, AD (2009). ERP correlates of processing native and nonnative language word stress in infants with different language outcomes. *Cortex*, **45**, 662–676.

Gervain, J, Macagno, F, Cogoi, S, Peña, M and Mehler, J (2008). The neonate brain detects speech structure. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, **105**, 14222–14227.

Gómez, DM, Berent, I, Benavides-Varela, S, Bion, RA, Cattarossi, L, Nespor, M and Mehler, J (2014). Language universals at birth. *Proceedings of the National Academy of Sciences*, **111**, 5837–5841.

Gonzales, K and Lotto, AJ (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, **24**, 2135–2142.

Greenberg, S, Carvey, H, Hitchcock, L and Chang, S (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics*, **31**, 465–485.

Grosjean, F (1982). *Life with Two Languages*. Cambridge, MA: Harvard University Press.

Hay, JSF and Diehl, RL (2007). Perception of rhythmic grouping: testing the iambic/trochaic law. *Perception of Psychophysics*, **69**, 113–122.

Hay, JF and Saffran, JR (2012). Rhythmical grouping biases constrain infant statistical learning. *Infancy*, **17**, 610–641.

Hayes, B (1995). *Metrical stress theory: principles and case studies*. Chicago & London: The University of Chicago Press.

Hayes, B (1985). Iambic and trochaic rhythm in stress rules. In Niepokuj M *et al.* (eds), *Proceedings of the Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*.

Höhle, B, Bijeljac-Babic, R, Herold, B, Weissenborn, J and Nazzi, T (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, **32**, 262–274.

Iversen, JR, Patel, AD and Ohgushi, K (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, **124**, 2263–2271.

Jusczyk, P and Bertoncini, J (1988). Viewing the development of speech perception as an innately guided learning process. *Language and Speech*, **31**, 217–238.

Kuhl, PK (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, **5**, 831.

LaCross, A (2015). Khalkha Mongolian speakers' vowel bias: L1 influences on the acquisition of non-adjacent vocalic dependencies. *Language, Cognition and Neuroscience*, **30**, 1033–1047.

Langus, A, Seyed-Allaei, S, Uysal, E, Pirmoradian, S, Marino, C, Asaadi, S, Eren, Ö, Toro, JM, Peña, M, Bion, RAH and Nespor, M (2016). Listening natively across perceptual domains?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 1127–1139.

Lerdahl, F and Jackendoff, R (1983). *Generative Theory of Tonal Music*, Cambridge: MIT Press.

Luk, G and Bialystok, E (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, **25**, 605–621.

Marian, V, Blumenfeld, HK and Kaushanskaya, M (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, **50**, 940–967.

Maye, J, Werker, JF and Gerken, L (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**, B101-B111.

Mehler, J, Dommergues, JY, Frauenfelder, U and Segui, J (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, **20**, 298–305.

Mersad, K and Nazzi, T (2011). Transitional probabilities and phonotactics in a hierarchical model of speech segmentation. *Memory and Cognition*, **39**, 1085–1093.

Michelas, A and D'Imperio, M (2010). Durational cues and prosodic phrasing in French: evidence for the intermediate phrase. *Speech Prosody* 2010, **100881**, 1–4.

Molnar, M, Lallier, M and Carreiras, M (2014). The amount of language exposure determines nonlinguistic tone grouping biases in infants from a bilingual environment. *Language Learning*, **64**, 45–64.

Moreton, E (2008). Analytic bias and phonological typology. *Phonology*, **25**, 83–127.

Narmour, E (1990). *The analysis and cognition of basic melodic structures*, Chicago, IL: University of Chicago Press.

Nazzi, T, Bertoncini, J and Mehler, J (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 756–766.

Nazzi, T, Iakimova, G, Bertoncini, J, Frédonie, S and Alcantara, C (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, **54**, 283–299.

Nespor, M, Shukla, M, Vijver, R. v. d, Avesani, C, Schraudolf, H and Donati, C (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e linguaggio*, **7**, 1–28.

New, B, Pallier, C, Ferrand L and Matos, R (2001). *Une base de données lexicales du français contemporain sur internet* [An online contemporary French lexical database]: LEXIQUE, Version 3.55, L'Année Psychologique, 101, 447–462. http://www.lexique.org. Accessed on 2010-11-11.

Peperkamp, S, Vendelin, I and Dupoux, E (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, **38**, 422–430.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rice, CC (1992). Binarity and ternarity in metrical theory: Parametric extensions. Unpublished doctoral dissertation, University of Texas, Austin.

Saffran, JR, Aslin, RN and Newport, EL (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.

Scrucca, L, Fop, M, Murphy, TB and Raftery, AE (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, **8**, 205–233.

Schwarz, GE (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.

Sebastián-Gallés, N, Echeverria, S and Bosch, L (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, **52**, 240–255.

Shi, R, Werker, JF and Morgan, JL (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, **72**, B11-B21.

Soares, C and Grosjean, F (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, **12**, 380–386.

Skoruppa, K, Pons, F, Christophe, A, Bosch, L, Dupoux, E, Sebastián-Gallés, N and Peperkamp, S (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, **12**, 914–919.

Skoruppa, K, Pons, F, Bosch, L, Christophe, A, Cabrol, D and Peperkamp, S (2013). The development of word stress processing in French and Spanish infants. *Language Learning and Development*, **9**, 88–104.

Spierings, M, Hubert, J and ten Cate, C (2017). Selective auditory grouping by zebra finches: testing the iambic–trochaic law. *Animal Cognition*, **20**, 665–675.

Sundara, M and Polka, L (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition*, **106**, 234–258.

Todd, N (1985). A model of expressive timing in tonal music. *Music Perception*, **3**, 33–58.

Toro, JM, Sebastián-Gallés, N and Mattys, SL (2009). The role of perceptual salience during the segmentation of connected speech. *European Journal of Cognitive Psychology*, **21**, 786–800.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press

Tsuji, S, Gomez, NG, Medina, V, Nazzi, T and Mazuka, R (2012). The labial–coronal effect revisited: Japanese adults say pata, but hear tapa. *Cognition*, **125**, 413–428.

Tyler, MD and Cutler, A (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, **126**, 367–376.

Venables, WN and Ripley, BD (2002). *Modern Applied Statistics with S. Fourth Edition*. New York: Springer.

Vouloumanos, A and Werker, JF (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, **10**, 159–164.

Vroomen, J, Tuomainen, J and de Gelder, B (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, **38**, 133–149.

Westfall, J (2015). PANGEA: Power analysis for general anova designs. Unpublished manuscript. Available at http://jakewestfall.org/publications/pangea.pdf.

Wickham, H (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer.

Woodrow, H (1909). A quantitative study of rhythm: The effect of variations in intensity, rate, and duration. *Archives of Psychology*, **14**, 1–66.

Woodrow, H (1951). Time perception. In Stevens SS (ed), *Handbook of Experimental Psychology*. New York: Wiley, pp. 1224–1236.

Yoshida, KA, Iversen, JR, Patel, AD, Mazuka, R, Nito, H, Gervain, J and Werker, JF (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition*, **115**, 356–361.

Winter, B (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv preprint arXiv*, **1308.5499**, 1–42.

https://doi.org/10.1017/S1366728920000140 Published online by Cambridge University Press