Effect of voicing on perceptual auditory laterality in Estonian and Norwegian native speakers

RENÉ WESTERHAUSEN
University of Oslo

NELE PÕLDVER and RICHARD NAAR
University of Tartu

DOMINIKA RADZIUN
University of Oslo

MARIA SILVIA KAAREP and KAIRI KREEGIPUU
University of Tartu

KENNETH HUGDAHL
University of Bergen and Haukeland University Hospital

PÄRTEL LIPPUS
Univerrsity of Tartu

KRISTIINA KOMPUS
University of Bergen and University of Tallinn

ADDRESS FOR CORRESPONDENCE
René Westerhausen, Research Group for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo, POB 1094, Blindern, 0317 Oslo, Norway. E-mail: rene.westerhausen@psykologi.uio.no

ABSTRACT
As a reliable and valid measures of perceptual auditory laterality, dichotic listening has been successfully applied in studies in many countries and languages. However, languages differ in the linguistic relevance of change in initial phoneme of words (e.g., for word identification). In the present cross-language study, we examine the effect of these differences on dichotic-listening task performance to establish how characteristics of one's native language affect the perception of nonnative phonological features. We compared 33 native speakers of Norwegian, a language characterized by a clear distinction between voiced and unvoiced initial plosive consonants, with 30 native speakers of Estonian, a language that has exclusively unvoiced initial phonemes. Using a free-report dichotic-listening paradigm utilizing pairs of voiced (/ba/, /da/, /ga/) and unvoiced (/pa/, /ta/, /ka/) stop-consonant vowels as stimulus material, the Norwegian native speakers were found to be more sensitive to the voicing of the initial

plosive than the Estonian group. "Voicing" explained 69% and 18% of the variance in the perceptual auditory laterality in the Norwegian and the Estonian sample, respectively. This indicates that experiential differences, likely during acquisition of the mother tongue in early development, permanently shape the sensitivity to the voicing contrast.

Keywords: brain asymmetry, dichotic listening, experiential differences, perceptual auditory laterality, voicing

The term *perceptual laterality* refers to systematic differences in the identification of stimuli presented to the left and right perceptual hemispaces (e.g., left and right visual half fields, left and right hand, and left and right ear) and are widely accepted to reflect underlying processing differences between the left and right brain hemispheres (e.g., Hellige, 1993). Dichotic listening is arguably the most frequently applied paradigm for the assessment of perceptual laterality in the auditory domain (Hugdahl, 2011). A typical trial of a dichotic-listening experiment consists of a simultaneous presentation of two different acoustic stimuli via headphones, whereby one of the stimuli is presented on the left and another one on the right sound channel (Bryden, 1988; Hugdahl, 2003). Asked to report the stimulus heard best after each trial, participants commonly identify and report more right-ear than left-ear stimuli when verbal stimuli, such as syllables or words, are used. This right-ear advantage is thought to reflect left hemispheric dominance for speech and language processing (for review see, e.g., Hiscock & Kinsbourne, 2011; Tervaniemi & Hugdahl, 2003; Toga & Thompson, 2003). One way to achieve a reliable and valid dichotic listening effect is to pair two sounds that are similar but not identical. That is, the spectrotemporal profiles of the two sounds should maximally overlap while at the same time the two stimuli need to fall into distinguishable perceptual categories (Wexler, 1988). In the most frequently used approaches, this is achieved by presenting verbal stimuli that only differ in the initial phoneme. For example, words that rhyme and differ only in the starting letter (e.g., *pin* and *bin*; see Fernandes, Smith, Logan, Crawley, & McAndrews, 2006; Wexler & Halwes, 1983) or consonant–vowel syllables with varying initial consonant but constant vowel (e.g., *pa* and *ba*; see Hugdahl & Andersson, 1986; Hugdahl et al., 2009) have been used routinely. Such initial-phoneme switch approaches have been successfully applied in demonstrating a right-ear advantage across many different languages, besides English (e.g., Arciuli, Rankine, & Monaghan, 2010) also Norwegian (e.g., Kompus et al., 2012), Swedish (Hugdahl & Andersson, 1986), Dutch (Van der Haegen, Westerhausen, Hugdahl, & Brysbaert, 2013), German (Westerhausen et al., 2006), Finnish (Takio et al., 2009), Spanish (Gadea et al., 2009), Italian (Brancucci et al., 2008), Turkish (Bayazıt, Öniz, Hahn, Güntürkün, & Özgören, 2009), or Japanese (Tanaka, Kanzaki, Yoshibayashi, Kamiya, & Sugishita, 1999), to name a few. At the same time, however, it is known that languages differ in the linguistic relevance of change in initial phoneme of words so that apparent differences in the magnitude of the advantage revealed in the direct comparison between different languages are difficult to interpret (Bless et al., 2015). For example, while Norwegian (as English) has a clear distinction between voiced (i.e., /b/, /d/, /g/) and unvoiced (i.e., /p/, /t/, /k/) initial plosive consonant phonemes (reflected, e.g., in the Norwegian minimal pair *gull* "gold" vs. *kull* "course"), Estonian does not have a comparable

Figure 1. (Color online) Illustration of difference in usage of initial stop-consonants in Estonian and Norwegian. Upper row, Estonian signs showing Estonian loan words or word stems starting with unvoiced plosives, that is, *kuld* ("gold"), *pangaautomaat* (literally "bank automate," i.e. "cash machine"), and *tantsustuudio* ("dancing studio"). Lower row, Norwegian signs containing the same words or word stems with voiced initial plosive, *gull* ("gold"), *bank*, and *dansestudio* ("dancing studio").

distinction. Rather, Estonian only has unvoiced plosive consonant phonemes within the standard language repertoire (Asu & Teras, 2009). The Estonian orthography uses the letters *g*, *d*, and *b* for a singleton unvoiced stop in the noninitial position, while the letters *k*, *t*, and *p* are used for their geminate counterparts (e.g., *kabi* /kapi/ "hoof" vs. *kapi* /kapːi/ "locker, sg. gen."). In the initial position, the consonants are always short, but in the orthography the letters *k*, *p*, and *t* are used. In loan words from other languages, the initial voiced stops are replaced by their unvoiced counterpart (see Figure 1). For example, the word *pank* is the Estonian word for *bank* (as loaned from the Germanic language family). In the newer loan words, often the letters *g*, *d*, and *b* at the beginning of words are remained in the spelling, but they are pronounced identically to their unvoiced counterpart as the phonemes /k/, /t/, and /p/, respectively. For example, the Estonian word *garaaž* "garage" is pronounced as /kɑrɑːʃ/, or the Estonian words *dušš* "shower" and *tušš* "Indian ink" both are pronounced identically as /tuʃː/. The phonological contrast between the consonants *g* and *k*, *d* and *t*, and *b* and *p* in initial position is neutralized, being relevant only for the differentiation of homonyms in spelling. In this, the contrast between voiced and unvoiced initial plosive consonant as usually used in dichotic-listening paradigms represents a phonologically artificial situation for native Estonian speakers. Thus, when compared to speakers of languages for which the identification of consonant voicing is important for word identification

(e.g., languages of the Germanic language families), it can be predicted that native Estonian speakers are less influenced by differences in voicing properties of initial plosive consonants also in a verbal dichotic-listening task. This would be in line with previous evidence showing that early experiential tuning to phonological features of one's native language (e.g., Kuhl et al., 2006) has a persisting effect on the perception of nonnative speech in adult life (see, e.g., Best, McRoberts, & Goodell, 2001), which could be reflected in differences in hemispheric specialization between languages (Bless et al., 2015).

The suggested difference in the sensitivity to the initial voicing can be systematically studied based on the phenomenon that unvoiced syllables are preferably reported in dichotic listening, regardless of to which ear they are presented (e.g., Andersson, Llera, Rimol, & Hugdahl, 2008; Arciuli et al., 2010; Berlin, Lowe-Bell, Cullen, Thompson, & Loovis, 1973; Gerber & Goldman, 1971; Rimol, Eichele, & Hugdahl, 2006; Sandmann et al., 2007; Voyer & Techentin, 2009). This so-called stimulus dominance (as originally suggested by Speaks, Niccum, Carney, & Johnson, 1981) has been shown both in English and Norwegian native speakers (e.g., Rimol et al., 2006; Voyer & Techentin, 2009). For example, in a Norwegian sample Rimol et al. (2006) found that dichotic pairings of voiced with voiced (VV) syllables (e.g., /da/–/ba/) and of unvoiced with unvoiced (UU) syllables (e.g., /ta/–/pa/) produce the typical right-ear preference. In pairs combining syllables of different voicing, the unvoiced syllable was shown to dominate the response pattern. That is, in presenting a voiced syllable to the left ear and an unvoiced to the right ear (VU), a right-ear advantage was found that was substantially increased compared to the UU and VV conditions. In addition, when presenting an unvoiced syllable to the left and a voiced to the right ear (UV), a significant left-ear advantage was found. However, to the best of our knowledge, the stimulus dominance for unvoiced syllables has not been examined in native Estonian speakers. Assuming that the relevance of initial consonant voicing is reduced for native Estonian as compared to Norwegian speakers, the difference in voicing between the two syllables in VU and UV trials should be less salient for Estonian speakers. Thus, it would be predicted that the stimulus dominance effect is weaker in native Estonian speakers.

The present cross-language study was designed to test the above prediction by comparing the voicing effect in native Norwegian and Estonian speakers using a consonant–vowel dichotic-listening paradigm. The direct comparison between native Norwegian and Estonian speakers allows examining how early native language experience affects (a) later perception of nonnative phonological features and (b) the hemispheric asymmetries supporting this phonological processing.

METHODS

Participants

In total 63 participants, 30 native Estonian and 33 native Norwegian speakers, took part in the study. The Estonian sample consisted of 15 male and 15 female participants recruited at the University of Tartu, Estonia, and had a mean age of

25.1 years ($SD = 5.0$ years). The Norwegian sample consisted of 15 male and 18 female participants recruited at the Universities of Bergen and Oslo, Norway, with a mean age of 24.3 ($SD = 3.6$). The age difference between the two groups was not significant ($t_{61} = 0.70$; $p = .49$, Cohen $d = 0.18$). Only right-handed participants were recruited, and handedness was verified with the Edinburgh Handedness Inventory (Oldfield, 1971). Audiometric screening was conducted to assure hearing acuity and acuity symmetry between the ears, by testing for left- and right-ear threshold for pure tones of 500, 1000, 1500, and 2000 Hz. Only participants with an average (across all frequencies) threshold ≤20 dB on each ear, and interaural threshold difference ≤10 dB went on to the dichotic-listening test. The study was approved by the Regional Ethical Committee of Northern Norway, for the Norwegian part, and the Research Ethics Committee of the University of Tartu, for the Estonian part. All participants gave written informed consent before participation.

Stimulus material

The six stop-consonants phonemes /b/, /d/, /g/, /p/, /t/, and /k/ were combined with the vowel /a/ to form the consonant–vowel syllables for the experiment. The syllables were recorded natural speech and spoken by a male native Estonian speaker in a neutral tone and with constant intensity, whereby the speaker was explicitly instructed to produce the (for Estonian phonologically unusual) voiced /ba/, /da/, and /ga/ syllables. The same Estonian syllables were used for both language groups to exclude the possibility that between-group differences in the stimulus material would systematically affect the results. The voice-onset time of the unvoiced syllables was 61 ms for /pa/, 78 ms for /ta/, and 73 ms for /ka/, and for the voiced syllables 15 ms for /ba/, 16 ms for /da/, and 19 ms for /ga/. Stimulus duration varied between 391 and 425 ms for unvoiced, and 344 and 374 ms for the voiced syllables. In this, the voice-onset time for the Estonian syllables was in a similar range as for previously used Norwegian recordings. For example, Rimol et al. (2006) report a voice-onset time for unvoiced syllables between 69 and 75 ms, and for voiced syllables between 25 and 31 ms. To further validate the use of Estonian syllables for a native Norwegian sample, a pilot test before the experiment proper showed that Norwegian speakers were able to perfectly identify these syllables. Of note, this is also reflected in the analysis of homonyms (see next paragraph) and indicates comparable high correct identification rates in both language samples. The final dichotic stimuli were created by combining syllables into pairs. The resulting stereo sound files were played such that one of the syllables was presented on the left and another one on the right sound channel. The two syllables of each pair were aligned to achieve simultaneous onset of the stop occlusion of the consonant segment in the two channels. The resulting 30 possible dichotic left-right channel combinations were classified into four voicing categories: (a) voiced syllable presented on the left and unvoiced syllable on the right channel (VU; e.g., /ba/–/pa/), (b) unvoiced syllable on the left and a voiced presented on the right (UV; e.g., /pa/–/ba/); (c) voiced syllable presented on both channels (VV; e.g., /ba/–/da/), and (d) unvoiced syllable presented on both channels (UU; e.g., /pa/–/ta/).

In addition, six homonymic pairs were created with the left and right channels consisting of the same stimulus (e.g., /ba/–/ba/ and /pa/–/pa/). The homonymic pairs served as control stimuli to check whether participants were able to correctly identify the stimuli. With a mean of 17.96 out of a maximum of 18 (99.77%) in the Estonian (29 out of 30 participants had all correct), and 17.89 out of 18 (99.38%) in the Norwegian sample (29 out of 33 all correct), participants of both groups showed a high level of correct homonym identification. Given the overall high accuracy in both groups (which prevented a parametric analysis) a 2 × 2 chi-square test was employed to compare the relative amount of participants who had identified all homonyms correctly or not ("all correct" and "not all correct") between the two language groups. The 2 × 2 chi-square test comparing the relative amount of participants who had identified "all correct" and "not all correct" between the two samples was not significant ($\chi^2 = 1.66$, $df = 1$; $p = .357$, Cramer $V = 0.16$). Thus, the homonym analysis indicates that the used stimulus material was suitable for both language groups.

Experimental procedure

The total set of 36 stimuli (i.e., dichotic and homonymic pairs) was presented three times as three experimental blocks. Thus, the experiment consisted of 108 trials of which 90 were dichotic presentations and 18 were homonyms. Of the 90 dichotic stimuli, 27 were instances of the UV and VU voicing category and 18 of the UU and VV voicing category, respectively. The order of the stimuli was pseudorandomized and followed the presentation order of the standardized Bergen Dichotic Listening Test (Hugdahl & Andersson, 1986; Hugdahl et al., 2009). Stimuli pairs were presented via headphones with a stimulus-onset asynchrony of 4.5 s between consecutive trials, and participants were instructed to report after each trial the one stimulus that was heard the best, whereby only one response was allowed per trial. Responses were collected via a keyboard (number pad), on which six separate response keys (numbers 1 to 6) were marked with the name of the six syllables. As each trial consists of two simultaneously presented stimuli, the participant's response could fall into one of three categories: (a) a correct identification of the left-ear stimulus, (b) a correct identification right-ear stimulus, or (c) a "false" response (i.e., reporting a stimulus that was not presented). The number of correctly identified left- and right-ear stimuli were used for further analyses. However, to account for differences in the number of presentations per voicing category, the percentage of correct left- and right-ear identifications for each of the four stimulus categories was calculated and used for statistical analyses.

The general experimental setup and procedure were identical at the three sites of data collection (Bergen, Oslo, and Tartu). The experiment took place in a sound-shielded room, the participants sat in comfortable chair, and the experiment was controlled via a personal computer. The same E-Prime script (Psychology Software Tools, Sharpsburg, PA) was used to run the experiment and collect the participants' responses. Nevertheless, differences in the exact equipment that was available at each site (i.e., computer build, keyboard, and headphone models) could not be avoided. However, because our effect of interest was an interaction in a within site (repeated measure) design and exact timing is not critical for the present experiment

(i.e., only accuracy data was used), we have no indications to believe that these equipment differences had any relevant effect on our results.

Statistical analysis

The analysis was set up as four-way (mixed) analysis of variance (ANOVA) with the between-subject factors language group (Estonian vs. Norwegian) and sex, and the within-subject factors ear (left vs. right) and voicing (four levels: UV, UU, VV, VU) using the percentage of correct syllable identification as dependent variable. The factor sex was included as previous studies indicate difference between male and female participants (Hirnstein, Westerhausen, Korsnes, & Hugdahl, 2013; Voyer, 2011). Because differences in the effect of voicing on the ear preference between Estonian and Norwegian were predicted, the three-way interaction of Ear × Voicing × Native Language was the effect of interest in the present analysis. The significant level was set to 5%, and significant effects were followed up with appropriate lower level post hoc analyses, that is, using ANOVAs and *t* tests. The effect size of main or interaction effects was calculated as proportion of explained variance (η^2). In cases where the difference between left- and right-ear response accuracy is important, the laterality index (LI) was used as an additional effect-size measure. The LI was defined as the difference between the number of correct left- and right-ear reports divided by the sum of both (and multiplied by 100 to obtain percentage scores). The statistical analysis was conducted using IBM SPSS (Version 22), and G*Power (Version 3.1.9) software was used for test power calculations.

RESULTS

The results of the four-way ANOVA are shown in Table 1. Post hoc analyses for the four effects that yielded statistical significance will be presented in here. The effect of interest, that is, the Ear × Voicing × Language interaction was significant ($\eta^2 = 0.14$) indicating a differential effect of voicing on the ear preference in the two language groups (see Figure 2). To further explore this effect, lower level post hoc ANOVAs were calculated from two perspectives. First, separate three-way ANOVAs (factors: ear, voicing, and sex) found a significant interaction of Voicing × Ear in both the Estonian, $F (3, 84) = 5.88$, $p = .001$, and the Norwegian, $F (3, 93) = 42.74$, $p < .001$, samples, whereby the effect size was with 69% explained variance substantially larger in the Norwegian than in the Estonian sample (18% explained variance). In the Estonian sample, post hoc pairwise comparisons showed a significant (all ps $< .003$) right-ear advantage in the VU (LI = 31.0%), VV (LI = 14.4%), and UU conditions (LI = 9.2%), while there was no significant difference in the UV condition (LI = 3.4%, $p = .54$). In the Norwegian sample, a significant right-ear advantage (all ps $< .004$) was detected in the VU (LI = 67.8%), VV (LI = 13.7%), and UU conditions (LI = 12.0%), while a significant left-ear advantage (LI = –30.4%) was detected in the UV condition ($p = .002$). Of note, both samples also showed a significant right-ear advantage as indicated by significant main effects of ear, Estonian: $F (1, 28) = 21.93$, $p < .001$; $\eta^2 = 0.45$, LI = 14.1%, and Norwegian: $F (1, 31) = 17.77$, $p < .001$; $\eta^2 = 0.17$;

Table 1. *Results of the four-way mixed analysis of variance with the between-subject factors language group (Estonian vs. Norwegian) and sex, and the within-subject factors ear and voicing*

Effect	F	df_{effect}	df_{error}	p	η^2
Ear	38.10	1	59	<.001	0.24*
Language	1.47	1	59	.229	<0.01
Sex	1.48	1	59	.228	<0.01
Voicing	124.08	3	177	<.001	0.11*
Ear × Language	0.05	1	59	.816	<0.01
Ear × Sex	0.01	1	59	.937	<0.01
Ear × Voicing	42.59	3	177	<.001	0.43*
Language × Sex	0.02	1	59	.898	<0.01
Language × Voicing	1.02	3	177	.386	<0.01
Sex × Voicing	2.49	3	177	.062	<0.01
Ear × Language × Sex	1.01	1	59	.320	<0.01
Ear × Voicing × Language	13.37	3	177	<.001	0.14*
Ear × Sex × Voicing	0.70	3	177	.556	<0.01
Language × Sex × Voicing	0.16	3	177	.926	<0.01
Ear × Language × Sex × Voicing	0.96	3	177	.415	<0.01

Note: η^2, effect size, explained variance.
*$p < .05$.
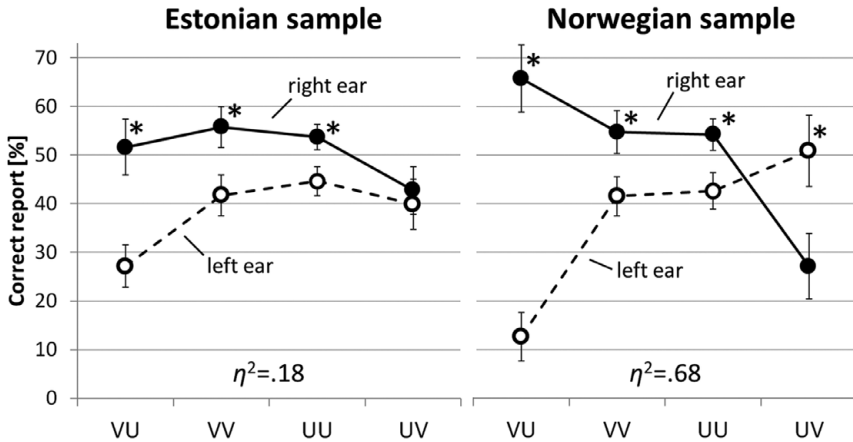

Figure 2. Voicing effect in Estonian and Norwegian sample. The graph shows the mean (± 95% confidence limits) percentage of correct left- and right-ear report per language group and for the four voicing conditions: VU, voiced syllable presented on the left and unvoiced syllable on the right auditory channel (e.g., /ba/–/pa/); VV, voiced syllable presented on both channels (e.g., /ba/–/da/); UU, unvoiced syllable presented on both channels (e.g., /pa/–/ta/); and UV, unvoiced syllable left and a voiced presented on the right (e.g., /pa/–/ba/). The effect size (η^2) quantifies the effect of voicing on the ear preference, and it represents the percentage of variance explained by the Voicing × Ear interaction in post hoc analyses of variance calculated separately for the Estonian and Norwegian subsamples (for details see text). Asterisks indicate significant pairwise comparisons between left- and right-ear correct responses (all ps < .004).

LI = 15.6%. In addition, in both samples the main effect of voicing was significant, Estonian: F (3, 84) = 62.71, $p < .001$; $\eta^2 = 0.21$, and Norwegian: F (3, 93) = 63.45, $p < .001$; $\eta^2 = 0.08$, and revealed a comparable pattern in post hoc tests: the percentage of correct answers was significantly higher (all $ps < .001$) in the conditions of equal voicing (UU, VV) compared to the mixed-voicing conditions (VU, UV), while there was neither a difference between UU and VV (all $ps > .30$) nor between the VU and UV conditions (all $ps > .07$).

In a second post hoc analysis of the three-way interaction, with the aim to identify which voicing conditions drive the above-indicated language group, we calculated separate post hoc ANOVAs (factors: ear, sex, and language) for the four voicing conditions. The interaction of Language × Ear, which is relevant for this purpose, was significant in the VU condition, F (1, 59) = 13.79, $p < .001$; $\eta^2 = 0.11$, where a stronger right-ear advantage was found in the Norwegian than in the Estonian group. The interaction was also significant in the UV condition, F (1, 59) = 9.88, $p = .003$; $\eta^2 = 0.36$, which was driven by a left-ear advantage in the Norwegian and a small (nonsignificant, see above) right-ear preference in the Estonian sample. However, the Language × Ear interaction was neither significant in the VV, F (1, 59) < 1, $p = .89$; $\eta^2 < 0.001$, nor in the UU condition, F (1, 59) < 1, $p < .57$; $\eta^2 = 0.01$; that is, here the magnitude of the ear advantage was comparable in both groups.

In addition to the three-way interaction, the four-way ANOVA revealed three significant effects (see Table 1 for test statistics). First, the main effect of ear was significant ($\eta^2 = 0.24$), indicating a right-ear advantage across all groups and conditions (LI = 14.8%). Second, the main effect of voicing yielded significance ($\eta^2 = 0.11$) with post hoc pairwise comparison showing the overall percentage of correct answers to be higher in the conditions of equal voicing (UU, VV) as compared to the mixed-voicing conditions (UV, VU; all $ps < .001$), while there was no difference between the UU and VV ($p = .27$) and between the UV and VU ($p = .26$) conditions, respectively. Third, the interaction of Ear × Voicing was significant ($\eta^2 = 0.43$). Post hoc analysis showed that this interaction is based on a significant (all $ps < .016$) right-ear advantage in the VU (LI = 49.4%), UU (LI = 14.1%), and VV conditions (LI = 10.6%), and a left-ear advantage in the UV condition (LI = –13.0%). No other main or interaction effects yielded significance (all $\eta^2 < .01$). Of importance, the main effect of language was not significant, indicating comparable overall performance levels in both groups.

DISCUSSION

The present analysis showed (a) that the interaction of Voicing × Ear was, with 69% compared to 18% explained variance, significantly stronger in the Norwegian than in the Estonian sample; and (b) that these sample difference were present in the mixed voicing (UV, VU) conditions but not in the equal voicing conditions (UU, VV). Together, these observations indicate that the stimulus dominance of unvoiced plosive consonant–vowel syllables is weaker in native Estonian than in native Norwegian speakers. Looking at the pattern of response across the four voicing conditions, the results of the present Norwegian sample replicate the findings

of several independent previous studies on Norwegian speakers (e.g., Andersson et al., 2008; Rimol et al., 2006; Sandmann et al., 2007): a moderate significant right-ear preference in the two conditions contrasting stimuli of the same voicing category, a substantially pronounced right-ear advantage in the VU condition, and a left-ear advantage in the UV condition. These findings were also in line with findings from English-speaking samples showing a comparable stimulus dominance effect of unvoiced syllables (e.g., Berlin et al., 1973; Gerber & Goldman, 1971; Voyer & Techentin, 2009). The present Estonian data, however, deviates from this pattern in two ways: in the VU condition the right-ear advantage was not as strongly accentuated relative to the same-voicing category conditions, and in the UV condition, no significant ear preference was found (cf. Figure 2). Thus, although it can be assumed that Estonian native speakers are being highly exposed to languages with contrastive voiced-unvoiced stops in adult everyday life (i.e., both Russian or English are widely used; see European Commission, 2012) and were able to almost perfectly identify unvoiced syllables when presented binaurally (homonym identification), their response to dichotic presented syllables was substantially less susceptible to voicing differences than it was for Norwegian speakers. As such, the observed group differences are likely linked to experiential differences during early development and acquisition of the mother tongue (for review, see Galle & McMurray, 2014). It has been shown that toward the end of the first year of life, the sensitivity toward native-language phonological contrasts increases while, at the same time, a decline in the sensitivity to nonnative phonological features can be observed (e.g., Kuhl et al., 2006). Related to the present findings, a Norwegian-speaking environment might have allowed the developing child to establish sensitivity to the initial unvoiced–voiced contrast, while it might not have been the case in an Estonian language environment where these contrasts do not exist. However, although this experiential tuning to one's native language takes place during infancy and early childhood, it has also been shown that the stimulus dominance for unvoiced syllables in dichotic listening is not fully developed until school age (Andersson et al., 2008; Westerhausen, Helland, Ofte, & Hugdahl, 2010) and may be associated to beginning of literacy education in school and the developing of phonological awareness (Ziegler & Goswami, 2005). Hence, it can be speculated that language experiences beyond infancy might contribute to the differences in the voicing effect between Estonian and Norwegian native speakers, while it remains for future developmental studies to test this hypothesis. Following the distinction between top-down attentional and bottom-up "hard-wired" processes in dichotic listening (Hiscock, Inch, & Kinsbourne, 1999, Hugdahl et al., 2009), it also remains to be determined on which stage of speech processing the reduced sensitivity to voicing features is manifested. Native Estonian speakers may pay less attention to the voicing features, or alternatively, not have developed the bottom-up neuronal sensitivity to process the contrast between voiced and unvoiced syllables. In both cases, differential implicit task-processing strategies might have been the consequence: native Estonians speakers might have to rely on acoustic stimulus characteristics for their response, while native Norwegians might utilize higher level phonetic processing.

Both the Norwegian and the Estonian sample showed a significant right-ear preference, and no difference in the magnitude of the right-ear advantage was found.

This was observed both as an average across the four voicing conditions as well as when only considering the (within) voicing category conditions UU and VV (which should not be affected by voicing differences). Thus, although Norwegian and Estonian speakers were differentially affected by the stimulus voicing, these effects appeared to be orthogonal to the underlying laterality effect as they were fully compensated for by averaging across all voicing conditions. Taking dichotic listening (right-) ear advantage to reflect underlying hemispheric specialization for speech and language processing (e.g., Hiscock & Kinsbourne, 2011; Tervaniemi & Hugdahl, 2003; Van der Haegen et al., 2013, Westerhausen, Kompus, & Hugdahl, 2014), the present study does not provides any indication for differences in hemispheric specialization between native Estonian and Norwegian speakers. The present study, however, does not replicate the finding of a previous study that reported reduced magnitude of the right-ear advantage in Estonian speakers as compared to native speakers of several other languages, including Norwegian, German, and English (Bless et al., 2015). At the same time, a post hoc power estimation for the present analysis showed sufficient test power (.81) for a replication (for a two-tailed t test at $\alpha = 0.05$), taking the empirical effect size of the difference between the Estonian and Norwegian laterality ($d = 0.73$) from the Bless et al. paper as a basis for the power estimation. A closer comparison of the two studies shows that although both used the same basic dichotic-listening paradigm, that is, both used a free-recall task instruction and plosive consonant–vowel syllables as stimulus material, the studies differed in some aspects that might have potentially contributed to the divergent findings, although not fully explain them. First, the Bless et al. (2015) study was based on data collection via a smartphone application and "crowd-sourced" participation, which, compared to the present laboratory experiment, is associated with reduced control over the testing environment and the technical equipment (e.g., which headphones were used) and composition of the study sample. Nevertheless, it has been previously demonstrated that the used smartphone application produces valid and reliable estimates of laterality (Bless et al., 2013), and it appears unlikely that the use of the smartphone application would introduce systematic performance differences between the two language groups. Second, while Bless et al. (2015) used stimuli for each language group that were spoken by a respective native speaker. The Norwegian sample was tested with syllables spoken by a native Norwegian speaker, and the Estonian sample was tested with syllables spoken by a native Estonian speaker. In the present study, both groups were tested with the very same Estonian syllables. One might argue that using the "nonnative" stimuli could have biased the performance of the present Norwegian sample and as such reduced possible group difference. However, the data does not support this interpretation. The Norwegian sample showed an almost perfect identification of the homonyms spoken by an Estonian speaker, which also did not differ from the performance of the Estonian sample. Furthermore, the here obtained mean laterality index of LI = 15.6% of the Norwegian sample tested with Estonian syllables was comparable in magnitude with the LI = 18.3% reported by Bless et al. (2015) or earlier studies on Norwegian samples (e.g., LI = 17.5% in Kompus et al., 2012) tested with Norwegian syllables. Thus, rather than being driven by the Norwegian sample, the difference between the two studies appears to be due to performance difference between the Estonian samples. The present

Estonian samples yielded a substantially higher overall LI (14.1%) than what was reported in the Bless et al. (2015) study (LI = 6.1%), although the very same stimulus material was used in both studies. Thus, excluding systematic effects of experimental setup or stimulus material, it appears that the observed inconsistency between studies with respect to laterality difference between Estonian and Norwegian samples are likely due to random sampling bias.

Present analyses also indicate that the percentage of correct stimulus identification (regardless of ear) is higher for the two conditions in which dichotic syllables of the same voicing category (UU, VV) are presented than for the two mixed voicing conditions (UV, VU). This is in line with previous studies showing that trials presenting pairs from the same voicing category are characterized by not only smaller number of errors (i.e., neither reporting the left- nor the right-ear stimulus correctly) but also faster response times for correct responses (Rimol et al., 2006; Westerhausen, Passow, & Kompus, 2013). This has been explained by referring to the fact that in dichotic pairs consisting of syllables from the same voicing category, the two auditory channels have a greater spectral and temporal overlap compared to pairs consisting of differentially voiced syllables (Brancucci et al., 2008). In brief, the high overlap of the same-voicing pairs increases the likelihood that the two dichotic stimuli are perceived as one "fused" stimulus (Cutting, 1976; Hiscock et al., 1999; Westerhausen et al., 2013), which can be easily identified and reported by the participant. However, the low overlap of the mixed-voicing pairs are less likely to fuse so that the participant are confronted with two competing and difficult to distinguish stimuli, making the task of reporting the one stimulus heard the best more cognitively demanding and error prone (Westerhausen et al., 2013). Thus, the here found lower overall correct identification in the UV and VU conditions compared to the UU and VV conditions can be seen as a result of the higher difficulty of response selection. Furthermore, the lack of a significant interaction of Voicing × Language further indicates that the difficulty of response selection across all conditions was comparable for both language groups.

Finally, the present findings also bear consequences for the interpretation of the stimulus dominance effect of unvoiced syllables in dichotic listening. Rimol et al. (2006; as well as others, e.g., Arciuli et al., 2010) have suggested that the dominance of unvoiced syllables can be explained by the mode of the stimulus presentation in dichotic listening. Within stimulus pairs of unequal voicing, which are synchronized to the consonant "occlusion," the vowel or voice onset is naturally delayed in the unvoiced relative to the voiced syllable. It was argued that this delay, in turn, would result in a "backward masking" effect making the unvoiced syllable more intelligible. However, if due to the relative timing of the acoustical features, the stimulus dominance effect should be independent of the participant's language background. Thus, the impact of native language on the magnitude of the stimulus dominance effect demonstrated in the present study supports the notion that it reflects an aspect of speech processing itself rather than of the mode of presentation.

Conclusions

Although there were no group differences in the magnitude of the right-ear advantage, the Norwegian native speakers were found to be more sensitive to the

voicing of the initial plosive than the Estonian group. Thus, the language background, likely due to early experiential tuning, shapes an individual's perceptual sensitivity for certain features of the native language. Taking behavioral laterality as a marker for underlying hemispheric differences, language background appears not to have an effect on functional hemispheric asymmetries for speech and language processing. Future comparative studies have to show whether the present findings, which were obtained by comparing languages of the Germanic and the Finno-Ugric language families, extend to other languages and language families. In such, the current study could be seen as a needed contribution to a research field that has been dominated by theories and data based on predominately English language studies and generalized to the world population at large although English is native language for a minority of people.

ACKNOWLEDGMENTS

REFERENCES

Andersson, M., Llera, J. E., Rimol, L. M., & Hugdahl, K. (2008). Using dichotic listening to study bottom-up and top-down processing in children and adults. *Child Neuropsychology, 14*, 470–479.

Arciuli, J., Rankine, T., & Monaghan, P. (2010). Auditory discrimination of voice-onset time and its relationship with reading ability. *Laterality, 15*, 343–360.

Asu, E. L., & Teras, P. (2009). Estonian. *Journal of the International Phonetic Association, 39*, 367–372.

Bayazıt, O., Öniz, A., Hahn, C., Güntürkün, O., & Özgören, M. (2009). Dichotic listening revisited: Trial-by-trial ERP analyses reveal intra- and interhemispheric differences. *Neuropsychologia, 47*, 536–545.

Berlin, C. I., Lowe-Bell, S. S., Cullen, J. K. Jr., Thompson, C. L., & Loovis, C. F. (1973). Dichotic speech perception: An interpretation of right-ear advantage and temporal offset effects. *Journal of the Acoustical Society of America, 53*, 699–709.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775–794.

Bless, J. J., Westerhausen, R., Arciuli, J., Kompus, K., Gudmundsen, M., & Hugdahl, K. (2013). "Right on all occasions?"—On the feasibility of laterality research using a smartphone dichotic listening application. *Frontiers in Psychology, 13*, 42.

Bless, J. J., Westerhausen, R., Torkildsen, J. V. K., Gudmundsen, M., Kompus, K., & Hugdahl, K. (2015). Laterality across languages: Results from a global dichotic listening study using a smartphone application. *Laterality, 20*, 434–452.

Brancucci, A., Penna, S. D., Babiloni, C., Vecchio, F., Capotosto, P., Rossi, D., . . . Rossini, P., M. (2008). Neuromagnetic functional coupling during dichotic listening of speech sounds. *Human Brain Mapping, 29*, 253–264.

Bryden, M. P. (1988). An overview of the dichotic listening procedure and its relation to cerebral organization. In K. Hugdahl (Ed.), *Handbook of dichotic listening: Theory, methods, and research* (pp. 1–44). Chichester: Wiley.

Cutting, J. E. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review, 83*, 114–140.

European Commission. (2012). *Europeans and their languages. Factsheet Estonia. Special Eurobarometer 386*. Retrieved from http://www.gesis.org/eurobarometer-data-service/publications/

Fernandes, M., Smith, M., Logan, W., Crawley, A., & McAndrews, M. (2006). Comparing language lateralization determined by dichotic listening and fMRI activation in frontal and temporal lobes in children with epilepsy. *Brain and Language, 96*, 106–114.

Gadea, M., Marti-Bonmatí, L., Arana, E., Espert, R., Salvador, A., & Casanova, B. (2009). Corpus callosum function in verbal dichotic listening: Inferences from a longitudinal follow-up of relapsing-remitting multiple sclerosis patients. *Brain and Language, 110*, 101–105.

Galle, M. E., & McMurray, B. (2014). The development of voicing categories: A quantitative review of over 40 years of infant speech perception research. *Psychonomic Bulletin & Review, 21*, 884–906.

Gerber, S. E., & Goldman, P. (1971). Ear preference for dichotically presented verbal stimuli as a function of report strategies. *Journal of the Acoustical Society of America, 49*, 1163–1168.

Hellige, J. B. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.

Hirnstein, M., Westerhausen, R., Korsnes, M. S., & Hugdahl, K. (2013). Sex differences in language asymmetry are age-dependent and small: A large-scale, consonant–vowel dichotic listening study with behavioral and fMRI data. *Cortex, 49*, 1910–1921.

Hiscock, M., Inch, R., & Kinsbourne, M. (1999). Allocation of attention in dichotic listening: Differential effects on the detection and localization of signals. *Neuropsychology, 13*, 404.

Hiscock, M., & Kinsbourne, M. (2011). Attention and the right-ear advantage: What is the connection? *Brain and Cognition, 76*, 263–275.

Hugdahl, K. (2003). Dichotic listening in the study of auditory laterality. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (pp. 441–475). Cambridge, MA: MIT Press.

Hugdahl, K. (2011). Fifty years of dichotic listening research—Still going and going and . . . *Brain and Cognition, 76*, 211–213.

Hugdahl, K., & Andersson, L. (1986). The "forced-attention paradigm" in dichotic listening to CV-syllables: A comparison between adults and children. *Cortex, 22*, 417–432.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., Laine, M., & Hämäläinen, H. (2009). Attention and cognitive control: Unfolding the dichotic listening story. *Scandinavian Journal of Psychology, 50*, 11–22.

Kompus, K., Specht, K., Ersland, L., Juvodden, H. T., van Wageningen, H., Hugdahl, K., & Westerhausen, R. (2012). A forced-attention dichotic listening fMRI study on 113 subjects. *Brain and Language, 121*, 240–247.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*, F13–F21.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97–113.

Rimol, L. M., Eichele, T., & Hugdahl, K. (2006). The effect of voice-onset-time on dichotic listening with consonant–vowel syllables. *Neuropsychologia, 44*, 191–196.

Sandmann, P., Eichele, T., Specht, K., Jäncke, L., Rimol, L. M., Nordby, H., & Hugdahl, K. (2007). Hemispheric asymmetries in the processing of temporal acoustic cues in consonant-vowel syllables. *Restorative Neurology and Neuroscience, 25*, 227–240.

Speaks, C., Niccum, N., Carney, E., & Johnson, C. (1981). Stimulus dominance in dichotic listening. *Journal of Speech, Language, and Hearing Research, 24*, 430–437.

Takio, F., Koivisto, M., Jokiranta, L., Rashid, F., Kallio, J., Tuominen, T., . . . Hämäläinen, H. (2009). The effect of age on attentional modulation in dichotic listening. *Developmental Neuropsychology, 34*, 225–239.

Tanaka, S., Kanzaki, R., Yoshibayashi, M., Kamiya, T., & Sugishita, M. (1999). Dichotic listening in patients with situs inversus: Brain asymmetry and situs asymmetry. *Neuropsychologia, 37*, 869–874.

Tervaniemi, M., & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews, 43*, 231–246.

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience, 4*, 37–48.

Van der Haegen, L., Westerhausen, R., Hugdahl, K., & Brysbaert, M. (2013). Speech dominance is a better predictor of functional brain asymmetry than handedness: A combined fMRI word generation and behavioral dichotic listening study. *Neuropsychologia, 51*, 91–97.

Voyer, D. (2011). Sex differences in dichotic listening. *Brain and Cognition, 76*, 245–255.

Voyer, D., & Techentin, C. (2009). Dichotic listening with consonant–vowel pairs: The role of place of articulation and stimulus dominance. *Journal of Phonetics, 37*, 162–172.

Westerhausen, R., Helland, T., Ofte, S., & Hugdahl, K. (2010). A longitudinal study of the effect of voicing on the dichotic listening ear advantage in boys and girls at age 5 to 8. *Developmental Neuropsychology, 35*, 752–761.

Westerhausen, R., Kompus, K., & Hugdahl, K. (2014). Mapping hemispheric symmetries, relative asymmetries, and absolute asymmetries underlying the auditory laterality effect. *NeuroImage, 84*, 962–970.

Westerhausen, R., Passow, S., & Kompus, K. (2013). Reactive cognitive-control processes in free-report consonant–vowel dichotic listening. *Brain and Cognition, 83*, 288–296.

Westerhausen, R., Woerner, W., Kreuder, F., Schweiger, E., Hugdahl, K., & Wittling, W. (2006). The role of the corpus callosum in dichotic listening: A combined morphological and diffusion tensor imaging study. *Neuropsychology, 20*, 272.

Wexler, B. E. (1988). Dichotic presentation as a method for single hemisphere stimulation studies. In K. Hugdahl (Ed.), *Handbook of dichotic listening: Theory, methods, and research* (pp. 85–115). Chichester: Wiley.

Wexler, B. E., & Halwes, T. (1983). Increasing the power of dichotic methods: The fused rhymed words test. *Neuropsychologia, 21*, 59–66.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3–29.