

Hemispheric specializations affect interhemispheric speech sound integration during duplex perception

Basil C. Preisig, and Matthias J. Sjerps

Citation: [The Journal of the Acoustical Society of America](#) **145**, EL190 (2019); doi: 10.1121/1.5092829

View online: <https://doi.org/10.1121/1.5092829>

View Table of Contents: <https://asa.scitation.org/toc/jas/145/3>

Published by the [Acoustical Society of America](#)

Hemispheric specializations affect interhemispheric speech sound integration during duplex perception

Basil C. Preisig^{a),b)} and Matthias J. Sjerps^{b)}

Donders Institute for Brain, Cognition and Behavior, Radboud University, P.O. Box 9101, Nijmegen, 6500 HB, The Netherlands
basilpreisig@gmx.ch, m.j.sjerps@gmail.com

Abstract: The present study investigated whether speech-related spectral information benefits from initially predominant right or left hemisphere processing. Normal hearing individuals categorized speech sounds composed of an ambiguous base (perceptually intermediate between /ga/ and /da/), presented to one ear, and a disambiguating low or high F_3 chirp presented to the other ear. Shorter response times were found when the chirp was presented to the left ear than to the right ear (inducing initially right-hemisphere chirp processing), but no between-ear differences in strength of overall integration. The results are in line with the assumptions of a right hemispheric dominance for spectral processing.

© 2019 Acoustical Society of America

[RS]

Date Received: January 11, 2019 **Date Accepted:** February 13, 2019

1. Introduction

Hemispheric asymmetries in speech and language processing have a long history, and the two most often reported specializations concern the general left hemisphere specialization for speech and language processing, and a right hemisphere specialization for spectral processing [(Poeppel, 2003; Zatorre and Belin, 2001); for a review see McGettigan and Scott (2012)]. This combination of findings raises the question of whether the right or left hemisphere is better equipped for processing the speech-related spectral information contained in formants. Here, we rely on duplex perception—the phenomenon that parts of the speech signal that are presented to different ears become integrated over time—to investigate whether speech-specific spectral information processing is more efficient with initial right or left hemisphere processing.

The contralateral pathways dominate initial auditory information processing and even suppress the ipsilateral ones in case of competing information (Kimura, 1967). Indeed the dominance of initial contralateral processing has been widely supported by functional imaging, neuropsychological, and psychoacoustic studies [(Kimura, 1967; Pollmann *et al.*, 2002; Studdert-Kennedy and Shankweiler, 1970), for a review see Hugdahl and Westerhausen (2016)]. The concept of contralateral transfer has played an important role in our understanding of the *Right Ear Advantage* (REA), a phenomenon typically observed in dichotic listening (DL) paradigms. In verbal DL experiments, two different speech sounds, commonly consonant-vowel (CV) syllables (e.g., /da/ and /ga/), are presented to one ear and the other ear. Participants typically report hearing the stimulus on the right ear (RE) more often, i.e., neglecting the left ear (LE) stimulus. Because of the dominance of contralateral information transfer, the REA in DL is generally interpreted as evidence for a left hemispheric dominance for speech perception (Hugdahl and Westerhausen, 2016).

Despite the dominance of the left hemisphere for speech and language processing, speech perception concerns an integration of information between both hemispheres. A clear example of such integration has been offered by the duplex perception phenomenon. In duplex perception paradigms, a short speech sound (e.g., a CV syllable) is split in different formant components that are presented separately to the two ears. Typically, a base stimulus (e.g., the first and second formant) is presented to one ear and a “chirp” (the third formant) is presented to the other. A single-formant chirp is usually perceived as non-speech sound when presented in isolation. However, when presented dichotically (i.e., with the remaining parts of the speech signal presented to

^{a)}Author to whom correspondence should be addressed.

^{b)}Also at: Max Planck Institute for Psycholinguistics, P.O. Box 310, Nijmegen, 6500 AH, The Netherlands.

the other ear), the chirp sound affects the perception of the “syllable” as a whole. This demonstrates that the two streams eventually become integrated (Liberman and Mattingly, 1989). It has been shown that integration in DL paradigms results from interhemispheric transfer across the corpus callosum (Steinmann *et al.*, 2014).

However, given the relative advantage of left hemisphere processing for speech sounds, along with the suggested right hemispheric dominance for spectral processing *per se*, it is unclear whether integration of the chirp sound would benefit more from initial left or right hemispheric processing. Some evidence exists that the processing advantages during duplex perception are different from DL of competing syllables (i.e., no REA for either LE or RE chirp presentation in healthy individuals) (Cranney *et al.*, 1989; Mathiak *et al.*, 2001). However, these studies presented a chirp transient to one ear with no signal at the other ear (i.e., there was no “competitor” formant). The lack of such a competitor is likely to result in a reduction of the strength of contralateral inhibition (Sparks and Geschwind, 1968), and hence the initial contralateral benefit. In the current study, we assessed whether a processing benefit would arise when a distinctive *F3* chirp is presented to one ear (with a frequency indicative of /ga/ or /da/), and a base sound that includes all formants, but for which the value of *F3* is ambiguous between /ga/ and /da/.

We hypothesized that right hemispheric specialization in spectral processing would increase the strength of the representation of the chirp and thereby facilitate binaural integration. This should result in more integration (i.e., more responses in accordance with the chirp *F3*) when the chirp is presented to the LE (and, hence, the base is presented to the RE). In addition, the facilitated processing should result in faster reaction times. Alternatively, given the overall dominance of left hemisphere processing for speech, one could also hypothesize that binaural integration might in fact be stronger when the chirp is presented to the RE (i.e., initially predominant left hemisphere processing). In line with the observation of the REA for speech, initial left hemisphere processing could result in privileged processing for the chirp when presented to the RE, with a representation that is more likely to be robustly integrated with the base sound. The current experiment assessed these two hypotheses.

2. Material and methods

2.1 Participants

Thirty-two right-handed volunteers [$M = 23.22$ yr, standard deviation (SD) = 3.60, 9 male] participated in the study. All participants had normal or corrected-to-normal visual acuity. The participants reported no history of neurological, psychiatric, or hearing disorders, and all had normal hearing (hearing levels of less than 25 dB tested at 250, 500, 750, 1000, 1500, 3000, and 4000 Hz, tested on both ears separately, including no interaural threshold differences above 5 dB), as assessed by pure tone audiometry (Maico MA30). All participants gave written informed consent prior to the experiment. Ethical approval to conduct this study was provided by the local ethics committee (CMO region Arnhem-Nijmegen). The present study was conducted in accordance with the principles of the latest version of the Declaration of Helsinki.

2.2 Stimuli

The stimuli created for the present study were based on stimuli reported in ten Oever and Sacks (2015), that were generously provided to us by those authors. Following ten Oever and Sacks, we manipulated the base stimulus (a recording of the stimulus /da/) which was down-sampled to 11 kHz and then morphed from a /da/ into a /ga/ stimulus by shifting the third formant in 17 equidistant steps from ~ 2.9 to ~ 2.5 kHz. This procedure was implemented using a source-filter separation procedure in Praat software (Boersma and Weenink, 2019). To generate the isolated chirps, the third formant (*F3*) was extracted from both endpoint stimuli (from /da/ and /ga/) by applying a bandpass filter with frequencies between 2100 and 3300 Hz. For a schematic representation of the stimuli see Fig. 1. For example stimuli, see supplementary material.¹

2.3 Procedure

Pretest. Because perceptual category boundaries can vary from person to person, a pretest was used to select participant-specific base stimuli for the main experiment. The goal of this manipulation was to create a stimulus that was highly sensitive to our chirp manipulation. Subjective category boundaries were estimated by assessing the individual psychometric curves (i.e., the point at which participants reported hearing a /da/ or a /ga/ in $\sim 50\%$ of the trials). Nine mono stimuli of the /da/ - /ga/ continuum were presented, each 16 times, in random order [Figs. 2(A) and 2(B)], resulting in a total of 144 trials. Participants who gave the “correct” response less than 50% of the

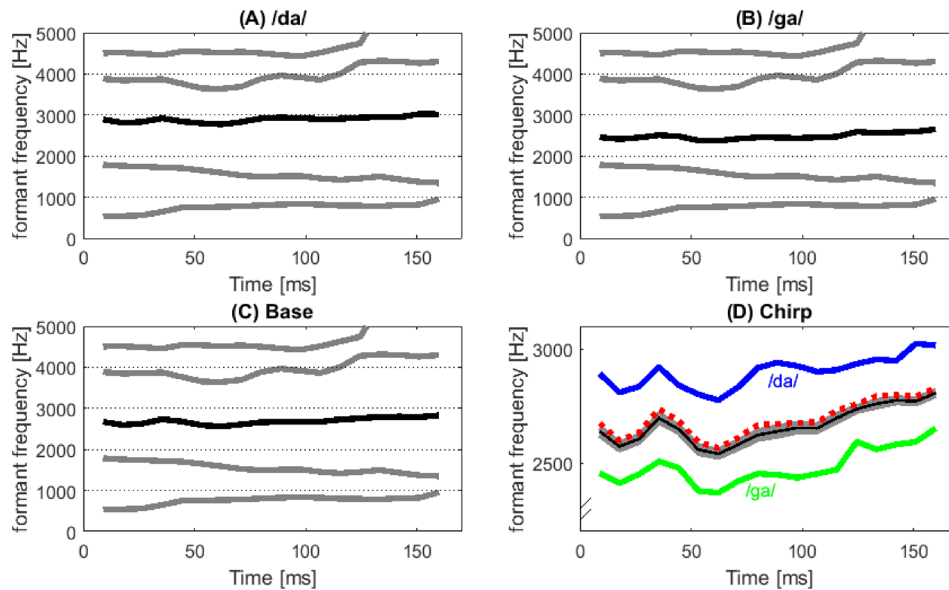


Fig. 1. (Color online) Schematic representation of the stimuli. (A) and (B) Stimulus properties of the /da/ and /ga/ endpoint stimuli. Only the third formant (black solid line) differs between the stimuli. (C) Exemplary ambiguous base stimulus, of which the third formant (black solid line) was adjusted to the individual category boundaries of the participant. (D) The chirp sounds consisting of only the third formant of the stimuli displayed in A (/da/) and B (/ga/). The dotted line represents the middle step of the continuum. The middle line represents the average category boundary (mean \pm SEM). Note that the y-axis scale in (D) differs from that in (A) to (C). SEM; Standard error of the mean.

time for the endpoint stimuli were excluded from further analyses (data from one participant were excluded based on this criterion). The average categorization responses of the remaining 31 participants are displayed in Fig. 2. To find the 50% crossover point, which represents an approximation of the individual category boundary, a binomial general linear model was fitted to the data of each participant individually. This model included the numerical independent variable Step (nine steps, ranging from /da/ to /ga/) and the dependent variable Response [0 = /da/; 1 = /ga/; using the fitglm function implemented in MATLAB statistics toolbox (Math Works, Natick, MA)]. The combination of the intercept and slope estimates were used to select the 50% crossover boundary for each participant [for individual examples see Fig. 2(B)]. The $F3$ frequency of the average individual category boundary was slightly lower than the middle step of the continuum [see Fig. 1(D)]. Thus, across our continuum, participants' responses were slightly biased toward /ga/ [$t(23) = 2.3782$, $p = 0.026$, one sample t -test].

Main experiment. An ambiguous base [see Fig. 1(C)] and a disambiguating $F3$ chirp [see Fig. 1(D)] were simultaneously presented to the two ears. The ambiguous

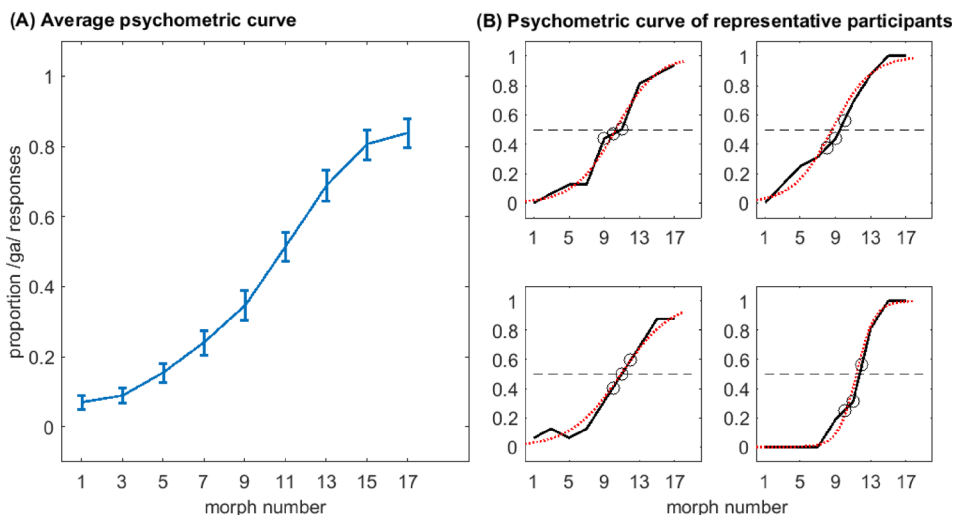


Fig. 2. (Color online) (A) Average proportion of /ga/ responses for the 31 participants (mean \pm SEM) who could reliably distinguish the response categories. (B) Psychometric curves of four representative participants. Black solid line, proportion of /ga/ responses as a function of morph number. Dotted line, quadratic fit. Circles, base morph steps selected for the main experiment. Error bars reflect the by-participant standard error of the mean (SEM).

base was based on participant-specific category boundaries, as determined in the pretest. For each participant, three stimuli were selected that fell close to the 50% crossover point of the individual psychometric curves [see Fig. 2(B)]. The chirp sound either supported a /da/ (high $F3 \sim 2.9$ kHz) or a /ga/ (low $F3 \sim 2.5$ kHz) interpretation. For the unambiguous trials (i.e., endpoint chirp + endpoint base stimuli), the value of the third formant was identical in both ears (both $F3$ values that supported a /da/ interpretation: ~ 2.9 kHz; or a /ga/ interpretation: ~ 2.5 kHz). In total, the participants completed 8 blocks, each consisting of 60 trials. Each block consisted of 48 chirp + ambiguous base trials and 12 chirp + endpoint base trials. The side of chirp presentation was alternated on every block. The order of the blocks was counterbalanced across participants.

The auditory stimuli were presented with an interstimulus interval of 3050 ms. Every stimulus was preceded by a fixation cross that was presented 610 ms before stimulus onset. After the fixation cross onset (1517 ms) the response options /ga/ and /da/ were presented one above and one below the fixation cross, falling within a visual angle of 9.43° . The participant indicated their response by pressing the corresponding respond button with their left index finger. It is important to note at this point that this study was part of a larger project assessing interhemispheric speech sound integration, for which an important part relied on functional Magnetic Resonance Imaging (fMRI). Since categorical decisions about speech sounds are typically dominated by ventral and dorsal left hemisphere processing, our participants were asked to make responses with their left (non-dominant) hand, which is controlled by the right hemisphere (see Hickok and Poeppel, 2007, for review). Although the current report does not include fMRI measurements, we aimed to implement minimal procedural changes between the different studies. Hence, here too, participants were asked to respond with their left (non-dominant) hand. We will further address to this aspect in the discussion (Sec. 4). Participants were instructed to perform as accurately and as fast as possible. The position of the response options was counterbalanced across participants. Auditory and visual stimulus presentation, as well as response recording, were controlled using Presentation[®] software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA).

2.4 Data analysis

In a first step, the reliability of the categorical judgments of individual participants on unambiguous endpoint trials during the main experiment was assessed. We tested whether the proportion of /ga/ responses differed between the stimulus types. For this purpose, a chi-square test was computed for every participant individually. Based on this criterion, the data of seven participants was excluded from further analyses because their classification accuracy did not significantly exceed chance level. The final dataset included data from 24 participants ($M = 23.00$ yr, $SD = 3.48$, 7 male). The average classification accuracy [%] for endpoint stimuli: /da/-endpoint ($M = 72.83$; $SD = 22.81$), /ga/-endpoint ($M = 77.78$; $SD = 22.34$).

Three dependent variables were analyzed, response category (/ga/;/da/ responses), integration of the chirp, i.e., responses consistent to the presented chirp, and reaction times. Statistical analyses were conducted in R (Version 3.3.3) using linear-mixed effect models as provided in the lme4 package (Bates *et al.*, 2014).

3. Results

3.1 The impact of isolated, contralateral $F3$ on the perception of an ambiguous base

First, it was tested whether participants' responses to ambiguous base stimuli were influenced by the presentation of the disambiguating chirp to the contralateral ear [Fig. 3(A)]. On average, participants gave 39.91% /ga/ responses for ambiguous bases combined with a /da/ supporting chirp (high $F3 \sim 2.9$ kHz) [$SD = 24.22$], and 70.75% /ga/ responses combined with a /ga/ supporting chirp (low $F3 \sim 2.5$ kHz) [$SD = 23.57$]. This difference was present in all participants and led to an average effect size (Cohen's d') of 1.474 [95% Confidence Interval (0.82, 2.13)].

To assess the influence of chirp $F3$ frequency on response category (0 = /da/; 1 = /ga/ response) logistic linear-mixed effect models (Baayen *et al.*, 2008; Jaeger, 2008) were fitted to the data. Independent variables included the fixed factors Chirp type (levels: high $F3 = -1$; low $F3 = 1$) and Base (three levels, i.e., three stimuli from the middle of the individual psychometric curves, coded as -1 , 0 , and 1), uncorrelated by-participant random intercepts and slopes were modeled as random effects. There were main effects of Chirp type [$\beta = 1.588$, Standard Error (SE) = 0.232; $z = 6.857$, $p < 0.001$], Base level ($\beta = 0.091$, SE = 0.029; $z = 3.165$, $p = 0.002$), and Intercept ($\beta = 0.445$, SE = 0.158; $z = 2.115$, $p = 0.034$), indicating that the $F3$ frequency of the

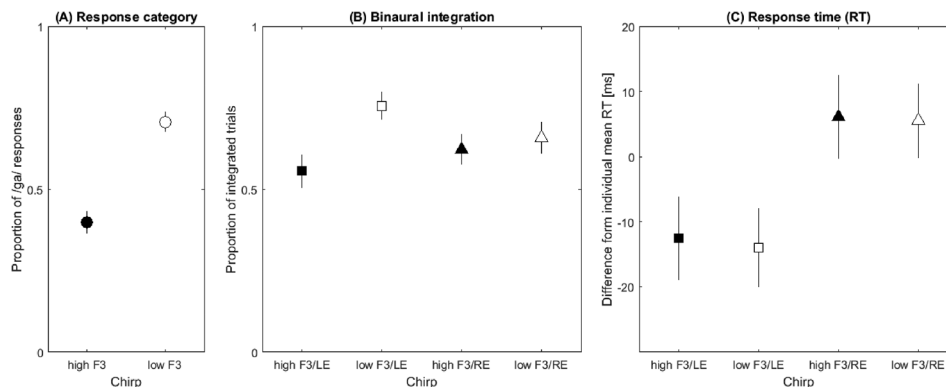


Fig. 3. (A) The proportion of /ga/ responses (mean \pm SEM across 24 participants) as a function of chirp (high F3, low F3). (B) Integration (mean \pm SEM across 24 participants), represented as the proportion of responses consistent with the chirp, and (C) response time, represented as the difference from the individual mean response time, as a function of side of chirp presentation [LE; RE] and chirp type (high F3; low F3).

chirp, but to some extent also the three levels of the base influenced syllable perception in the expected—integrative direction. There was no evidence for an interaction of Chirp type \times Base level ($\beta = -0.003$; SE = 0.058, $z = -0.047$, $p = 0.962$). These results thus indicate that the chirp was reliably integrated with the base sounds.

3.2 Stimulus laterality effect

Having established the occurrence of integration, we next assessed the influence of laterality and chirp F3 frequency on integration (0 = no integration, i.e., response inconsistent with the chirp; 1 = chirp integration, i.e., response consistent with the presented chirp). A logistic linear mixed effect model was fitted including the fixed factors Chirp type and Side of chirp presentation (LE = -1; RE = 1), uncorrelated by-participant random intercepts and slopes were modeled as random effects. The analysis revealed a main effect for Chirp type ($\beta = 0.753$, SE = 0.334; $z = 2.252$, $p = 0.024$), reflecting higher integration scores for low F3 (ga-chirp) presentation, but, critically, no effect for Side of chirp presentation ($\beta = -0.043$, SE = 0.088; $z = -0.491$, $p = 0.623$), and no interaction Chirp type \times Side of chirp presentation ($\beta = -1.055$, SE = 0.735; $z = -1.436$, $p = 0.151$). Overall, average integration was higher when the /ga/ chirp was presented ($M = 70.75\%$; SD = 23.57%) compared to /da/ chirp presentation ($M = 58.92\%$; SD = 24.66%) [Fig. 3(B)]. This result may reflect an overall tendency of the participants to give more /ga/ responses.

To test the impact of the variables Side of presentation and Chirp type on processing speed a linear mixed effect model including these factors was calculated for the dependent variable reaction time. Uncorrelated by-participant random intercepts and slopes were modeled as random effects. The model revealed a significant main effect of Side of presentation ($\beta = 21.139$, SE = 6.725; $t = 3.143$, $p = 0.002$), but no effect for Chirp type ($\beta = -4.816$, SE = 5.817; $t = -0.828$, $p = 0.408$) and no interaction Chirp type \times Side of presentation ($\beta = -9.078$, SE = 7.137; $t = -1.272$, $p = 0.203$). Reaction times were faster if the chirp was presented to the LE than if presented to the RE [Fig. 3(C)].

4. Discussion

The present study aimed to investigate whether speech-related spectral information processing benefits from a predominant initial left or right hemisphere processing. Here, we tested hemispheric asymmetries during binaural integration of CV syllables. Syllables were split in an ambiguous base (perceptually intermediate between /ga/ and /da/), which was presented to one ear, and a disambiguating F3 chirp to the other ear. We found no evidence that the side of F3 chirp presentation reliably influenced participants' ultimate perceptual decisions. However, we did find an effect on reaction times, suggesting faster processing of spectral information when the chirp is presented to the LE (and base is presented to the RE).

First, we established that the presentation of a disambiguating formant chirp showed the expected effect on response category (on average 60.09% /da/ responses on high F3 and 70.95% /ga/ responses on low F3 chirps). This result indicates that the chirp was reliably integrated with the base. However, we found no impact of ear of chirp presentation on the likelihood of integration. This finding corresponds with previous studies showing no ear advantages during duplex perception (Mathiak *et al.*, 2001; Rand, 1974). However, an important difference with those studies was that in their case the acoustic targets (formant chirp or formant transient) were only presented

unilaterally. For the present study, the base sounds included a perceptually intermediate $F3$ competitor. This approach is likely to reduce the strength of initial ipsilateral information transfer. In addition, previous studies used synthesized speech stimuli only. However, the current findings demonstrate that these differences in stimulus properties did not lead to fundamentally different behavioral outcomes.

Despite the lack of a laterality effect on ultimate decisions, we did find shorter response times when the $F3$ chirp was presented to the LE. We interpret this finding as resulting from a processing advantage of spectral information in the right hemisphere, as advocated, for example, in the asymmetric sampling theory (AST) (Poeppel, 2003; see Zatorre and Belin, 2001 for a similar framework). The AST suggests that auditory speech signals get processed asymmetrically in the time domain: left auditory areas preferentially extract information from short (~ 20 – 40 ms) temporal integration windows. The right hemisphere homologues preferentially extract information from longer (~ 150 – 250 ms) integration windows (Poeppel, 2003), which makes them better suited for slow-changing spectral information processing. Note that our chirp stimuli differed from each other across a 160 ms window [Fig. 1(D)], which means that within the AST framework they would be considered slow spectral changes. Hence, initial right hemisphere processing (following LE presentation) might be faster because the right hemisphere is assumed to be more efficient for processing these properties of sound (Poeppel, 2003; Zatorre and Belin, 2001). While it should be acknowledged that AST has been challenged [see McGettigan and Scott (2012), for review], there is continuing support for the more general idea that right hemisphere processing may be especially efficient for spectral information (e.g., see Bouton *et al.*, 2018 for recent evidence from brain lesions and intracranial recordings). Our findings thus partly align with this framework insofar as they demonstrate a processing advantage when the chirp is presented to the LE and the base presented to the RE, respectively.

However, one potential concern with the findings reported here is related to the fact that our participants responded with their left index finger; as outlined under Sec. 2.3 this decision was based on design restrictions for subsequent fMRI studies not reported here. Given this design, a chirp to the RE would be generally “disadvantaged,” because motor preparation would then require the additional transfer step over the corpus callosum from the receiving left hemisphere to the right hemisphere that controls the movement of the responding (left) hand. In other words, a stimulus that is presented on the same ear as the responding hand could, in principle, always be privileged. If this interpretation would hold, this would provide an alternative explanation for our findings. Importantly, however, previous DL research has provided robust evidence against this assumption. For example, in a large study of interhemispheric speech processing [Jäncke (2002), and for a subgroup of right handed participants, like our participants], it was found that RE stimuli are responded to faster overall, and that right hand responses were faster overall. But it was not the case that responding with the hand that was on the same side of the stimulus was faster than cases in which they mismatched [see Fig. 4 of Jäncke (2002)]. Moreover, for the left hand response condition, fastest response times were reported for RE stimuli (i.e., unlike our findings). This observation suggests that our observation of a LE advantage for chirp stimuli is specific to the nature of the stimuli used here. That is, a LE advantage is only observed when it concerns isolated spectral information processing.

The observation of shorter response times for the $F3$ chirps presented to the LE is insofar surprising as previous DL experiments have consistently found a REA for consonants, corresponding to preferential left-hemisphere processing for short consonant cues such as formant transitions or voice onset times. (Cutting, 1974; Schwartz and Tallal, 1980; Studdert-Kennedy and Shankweiler, 1970). However, previous evidence suggests that the processing advantages during duplex perception could be different from DL of syllables (i.e., no REA for either LE or RE chirp presentation in healthy individuals) (Cranney *et al.*, 1989; Mathiak *et al.*, 2001). Our interpretation is that this is due to a processing advantage of initial right hemisphere processing of the chirp that biases participants’ perception to one or the other stimulus percept.

An important additional consideration is that while we observe a LE processing advantage for chirp stimuli, this could also be interpreted as a REA for the base stimulus (since these are always presented contralaterally in the duplex perception paradigm). Indeed, the majority of phonetic information is provided by the ambiguous base. Thus, it is possible that the observed response time difference could, to some extent, also reflect a REA for the perception of the ambiguous base in terms of faster access to phonetic representations in the left hemisphere (although note that without the chirp information, the base remains ambiguous). Furthermore, it cannot be ruled out that the interhemispheric transfer of the richer phonetic information, as it is

provided by the base, is less efficient. Hence, while we argue that the findings presented here are most readily interpreted within the framework of a right hemisphere advantage for spectral processing, important insights can be gained from further studies investigating interhemispheric processing of speech.

5. Conclusion

We did not find conclusive evidence supporting the notion that laterality of initial chirp sound processing would influence participants' ultimate perceptual decisions. However, we did observe that laterality of chirp presentation affects the processing speed of binaural integration of speech. The latter observation supports the notion of right hemispheric specialization for spectral processing as formulated in theories such as AST.

Acknowledgments

This work was supported by the Swiss National Science Foundation (Grant Nos. P2BEP3_168728 and PP00P1_163726) and Janggen-Pöhn-Stiftung. The authors would like to thank Lenno Ruijters, Margot Mangnus, and Miriam Greidanus Romanelli for their assistance.

References and links

¹See supplementary material at <https://doi.org/10.1121/1.5092829> for exemplary stimuli.

- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). "Mixed-effects modeling with crossed random effects for subjects and items," *J. Mem. Lang.* **59**, 390–412.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). "Fitting linear mixed-effects models using lme4," June 23, 2014, available at <http://arxiv.org/abs/1406.5823>.
- Boersma, P., and Weenink, D. (2019). Praat: Doing Phonetics by Computer (Computer program), <http://www.praat.org/> (Last viewed February 17, 2019).
- Bouton, S., Chambon, V., Tyrand, R., Guggisberg, A. G., Seeck, M., Karkar, S., van de Ville, D., and Giraud, A. L. (2018). "Focal versus distributed temporal cortex activity for speech sound category assignment," *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1299–E1308.
- Cranney, J., Fowler, C. A., and Musiek, F. (1989). "Duplex perception: Some initial findings concerning its neural basis," *Brain Cogn.* **9**, 48–56.
- Cutting, J. E. (1974). "Two left-hemisphere mechanisms in speech perception," *Percept. Psychophys.* **16**, 601–612.
- Hickok, G., and Poeppel, D. (2007). "The cortical organization of speech processing," *Nat. Rev. Neurosci.* **8**(5), 393–402.
- Hugdahl, K., and Westerhausen, R. (2016). "Speech processing asymmetry revealed by dichotic listening and functional brain imaging," *Neuropsychologia* **93**, 466–481.
- Jaeger, F. (2008). "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models," *J. Mem. Lang.* **59**, 434–446.
- Jäncke, L. (2002). "Does 'callosal relay' explain ear advantage in dichotic monitoring?," *Lateral. Asymmetries Body Brain Cogn.* **7**, 309–320.
- Kimura, D. (1967). "Functional asymmetry of the brain in dichotic listening," *Cortex* **3**, 163–178.
- Liberman, A. M., and Mattingly, I. G. (1989). "A specialization for speech perception," *Science* **243**, 489–494.
- Mathiak, K., Hertrich, I., Lutzenberger, W., and Ackermann, H. (2001). "Neural correlates of duplex perception: A whole-head magnetencephalography study," *NeuroReport* **12**, 501–506.
- McGettigan, C., and Scott, S. K. (2012). "Cortical asymmetries in speech perception: What's wrong, what's right and what's left?," *Trends Cogn. Sci.* **16**, 269–276.
- Poeppel, D. (2003). "The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time,'" *Speech Commun.* **41**, 245–255.
- Pollmann, S., Maertens, M., von Cramon, D. Y., Lepsien, J., and Hugdahl, K. (2002). "Dichotic listening in patients with splenial and nonsplenial callosal lesions," *Neuropsychology* **16**, 56–64.
- Rand, T. C. (1974). "Dichotic release from masking for speech," *J. Acoust. Soc. Am.* **55**, 678–680.
- Schwartz, J., and Tallal, P. (1980). "Rate of acoustic change may underlie hemispheric specialization for speech perception," *Science* **207**, 1380–1381.
- Sparks, R., and Geschwind, N. (1968). "Dichotic listening in man after section of neocortical commissures," *Cortex* **4**, 3–16.
- Steinmann, S., Leicht, G., Ertl, M., Andreou, C., Polomac, N., Westerhausen, R., Friederici, A. D., and Mulert, C. (2014). "Conscious auditory perception related to long-range synchrony of gamma oscillations," *NeuroImage* **100**, 435–443.
- Studdert-Kennedy, M., and Shankweiler, D. (1970). "Hemispheric specialization for speech perception," *J. Acoust. Soc. Am.* **48**, 579–594.
- ten Oever, S., and Sack, A. T. (2015). "Oscillatory phase shapes syllable perception," *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15833–15837.
- Zatorre, R. J., and Belin, P. (2001). "Spectral and temporal processing in human auditory cortex," *Cereb. Cortex* **11**, 946–953.