Brain and Cognition 76 (2011) 233–238

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

# Manipulation of voice onset time during dichotic listening

Joanne Arciuli *

*Faculty of Health Sciences, University of Sydney, Australia*

ARTICLE INFO

*Article history:*
Available online 12 February 2011

*Keywords:*
Dichotic listening
DL
VOT
Attention
Lateralisation
Speech perception
Reading ability

ABSTRACT

The manipulation of voice onset time (VOT) during dichotic listening has provided novel insights regarding brain function. To date, the most common design is the utilisation of four VOT conditions: short–long pairs (SL), where a CV syllable with a short VOT is presented to the left ear and a CV syllable with a long VOT is presented to the right ear as well as long–short (LS), short–short (SS) and long–long (LL) pairs. Rimol, Eichele, and Hugdahl (2006) first reported that in healthy adults SL pairs elicit the largest REA while, in fact, LS pairs elicit a significant left ear advantage (LEA). This VOT effect was replicated by Sandmann et al. (2007). A study of children aged 5–8 years of age has shown a developmental trajectory whereby long VOTs gradually start to dominate over short VOTs when LS pairs are being presented under dichotic conditions (Westerhausen, Helland, Ofte, & Hugdahl, 2010). Two studies have investigated attentional modulation of the VOT effect in children and adults. The converging evidence from these studies shows that at around 9 years of age children lack the adult-like cognitive flexibility required to exert top-down control over stimulus-driven bottom-up processes (Andersson, Llera, Rimol, & Hugdahl, 2008; Arciuli, Rankine, & Monaghan, 2010). Arciuli et al. further demonstrated that this kind of cognitive flexibility is a predictor of proficiency with complex tasks such as reading. A review of each of these studies, the possible mechanisms underlying the VOT effect and directions for future research are discussed.


## 1. Introduction

During dichotic listening (DL) participants are exposed to competing stimuli presented to the right and left ear. The structural hypothesis of DL first proposed by Kimura (1967) suggests that the effects observed during DL reflect brain organisation.[1] Monaural input to the either the left or the right ear is processed by both hemispheres due to the presence of contralateral pathways and ipsilateral pathways; however, there is an advantage of contralateral over ipsilateral pathways due to the former being more numerous and more rapidly conducting. The simultaneous presentation of auditory stimuli appears to overload the perceptual system resulting in temporary isolation of the hemispheres. This isolation comes about primarily through suppression of the ipsilateral pathways (e.g., Wexler, 1988). Accordingly, during DL, stimuli presented to the right ear are directed to the temporal lobe of the left hemisphere while stimuli presented to the left ear are directed to the temporal lobe of the right hemisphere. Under dichotic conditions, auditory input may still reach ipsilateral areas via transfer across the corpus callosum; although, such input is delayed and depleted (Pollmann, Maertens, von Cramon, Lepsein, & Hugdahl, 2002). Techniques including PET, fMRI and MEG (e.g., Della Penna et al., 2007; Hugdahl et al., 1999; and Thomsen, Rimol, Ersland, & Hugdahl, 2004, respectively), electrophysiological measurements (e.g., Eichele, Nordby, Rimol, & Hugdahl, 2005) and the Wada-test (e.g., Hugdahl, Carlsson, Uverbrant, & Lundervold, 1997) have demonstrated that DL is an effective measure of hemispheric processing.

Whilst DL is not used exclusively for the examination of hemispheric processing of language, it is perhaps the most commonly used technique in the study of language laterality (O'Leary, 2003). The most often reported finding in the DL literature is that healthy, right-handed individuals exhibit a right ear advantage (REA) for speech. Although there is now clear evidence that both hemispheres are involved in speech perception, contrary to the classical Wernicke–Geschwind model, the left temporal lobe may be specialised for the processing of at least some aspects of linguistic material. In terms of the structural hypothesis of DL it has been shown that there is particularly strong suppression of ipsilateral pathways in the left hemisphere when speech is being presented under dichotic conditions (Della Penna et al., 2007). Thus, the REA for speech is considered to be the result of stimulus-driven bottom-up processing. This paper is a review of a relatively recent development in the DL literature concerning the presence of a left ear advantage (LEA) for linguistic stimuli that exhibit a certain type of voicing contrast. The methodology used to obtain such an effect has been variously described as "a more powerful determinant of

* Address: Faculty of Health Sciences, University of Sydney, P.O. Box 170, Lidcombe 1825, Australia.
*E-mail address:* joanne.arciuli@sydney.edu.au

[1] See Kinsbourne (1970) for an alternative interpretation.

0278-2626/$ - see front matter Crown Copyright © 2011 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2011.01.007

DL performance" (Rimol et al., 2006, p. 194), "more sensitive in detecting developmental changes" (Westerhausen et al., 2010, p. 759), and "a new way to study aspects of cognitive development" (Andersson et al., 2008, p. 477).

## 2. The manipulation of voicing during DL

A range of phonological features have been investigated in DL experiments. Segmental phonology has more often been the focus than suprasegmental phonology (although see Arciuli & Slowiaczek, 2007, for a recent investigation of the effects of lexical stress during DL). Liquids, fricatives and vowels show a reduced REA compared to stop consonants (Darwin, 1971; Haggard, 1971; Hugdahl & Andersson, 1984). Larger REAs have been found using initial rather than final consonants (Bryden, 1988). Some of the most robust DL effects have been demonstrated using CV syllables where initial stop consonants are paired with a common short vowel (Hugdahl, 1995). A small number of studies have investigated specific characteristics of stop consonants during DL (e.g., Voyer & Techentin, 2009 examined the role of place of articulation in CV pairs).

Of direct relevance to the current review there is renewed interest in the manipulation of voice onset time (VOT) in CV syllables during DL (see the work of Repp for earlier studies in this area, for example, Repp, 1977, 1978).[2] A stop consonant is realised by an occlusion of air-flow followed by a release of energy. VOT refers to the length of the interval between the release of the consonant and the onset of vocal cord vibration. Syllable initial voiced consonants (/b/) have short VOTs whereas their unvoiced counterparts (/p/) have long VOTs. VOT provides linguistically meaningful contrasts; for example, in distinguishing between 'pat' and 'bat'. Thus, the identification of voicing is important for word identification and comprehension processes.

Rimol et al. (2006) systematically manipulated VOT during DL in their study of 89 healthy right-handed adults (median age 25 years, 45 females). Stimuli included six stop consonants paired with the vowel /a/ within CV syllables. Three of the consonants had short VOTs between 25 and 31 ms ('ba', 'da', 'ga') and three consonants had long VOTs of between 69–75 ms ('pa', 'ta', 'ka'). There were four VOT conditions: short–long pairs (SL), where a CV syllable with a short VOT is presented to the left ear and a CV syllable with a long VOT is presented to the right ear as well as long–short (LS), short–short (SS) and long–long (LL) pairs. There were 30 different pairings of syllables, each presented a total of nine times. See Table 1 for syllable pairings in each of the four VOT conditions.

The researchers excluded participants that did not show an overall REA. They then conducted two types of analyses: (1) comparison of the percentage of correct reports in the right ear versus the left ear, and (2) comparison of a Laterality Index which is positive for a REA or negative for a LEA [(right ear correct reports−left ear correct reports)/total number of correct reports] against the value of zero. The largest REA was elicited when participants listened to SL pairs although SL, SS and LL pairs all elicited a REA (there was a significantly higher percentage of correct right ear versus left ear reports for SL, SS and LL pairs and the Laterality Index was significantly larger than zero for the SL, SS and LL pairs). The key finding was that a LEA emerged when participants were presented with LS pairs (there was a significantly higher percentage of left ear versus right ear reports and the Laterality Index was significantly less than zero for the LS pairs). Thus, for syllable pairs that exhibited voicing contrasts (SL and LS pairs), long VOT

**Table 1**
CV syllable pairs four VOT conditions (similar to the tables reported by Rimol et al. (2006) and Arciuli et al. (2010)).

| SL | LS | SS | LL |
|---|---|---|---|
| ba–ka | pa–ba | ba–da | pa–ka |
| ba–pa | pa–da | ba–ga | pa–ta |
| ba–ta | pa–ga | da–ba | ta–ka |
| da–ka | ta–ba | da–ga | ta–pa |
| da–pa | ta–da | ga–ba | ka–pa |
| da–ta | ta–ga | ga–da | ka–ta |
| ga–ka | ka–ba | | |
| ga–pa | ka–da | | |
| ga–ta | ka–ga | | |

syllables elicited a processing advantage regardless of the ear to which they were presented, even in participants who showed an overall REA. Further analyses were conducted excluding trials where syllable pairs differed both in place of articulation and VOT. The results did not change. Remember, the researchers excluded participants that did not show an overall REA. In doing so, they probably lessened the chances of observing any kind of LEA.

Rimol et al. interpreted their results in terms of different (and competing) influences, namely, two bottom-up, stimulus-driven factors. The first relates to a right ear advantage for linguistic stimuli and the second to advantaged processing of stop consonants that exhibit long VOTs (that can elicit either a REA for an SL pair, or an LEA for an LS pair). Rimol et al. did not interpret their findings in light of whether the right or the left hemisphere might be specialised for the processing of voicing *per se*. Rather, they were focussed on the mechanisms that are responsible for the brain resolving dichotically presented syllable pairs in favour of stop consonants with long VOTs. They discussed the possibility that syllables exhibiting long VOTs may require less temporal precision during processing thereby providing a more stable perceptual trace and the possibility that syllables with a long VOT may provide a backward masking of syllables with a short VOT. These possibilities seem more likely than one based on the premise that the right and left hemispheres are differentially specialised for processing longer and shorter temporal events, respectively (e.g., the Asymmetric Sampling in Time model, AST, described by Poeppel (2003)). The LEA effect for LS pairs might be taken as support for the hypothesis that the right hemisphere is better suited to the processing of long acoustic events (indeed, there is substantial evidence for this in the literature) but it must be noted that SL pairs elicited a REA that was significantly greater than that elicited by SS or LL pairs. Whatever the underlying mechanism, Rimol et al.'s conclusion that VOT manipulation during DL is an especially sensitive tool in the investigation of competing stimulus-driven processes has sparked considerable interest.

A study conducted by Sandmann et al. (2007) sought to replicate the VOT effect behaviourally and examine its time-course electrophysiologically. Forty-six healthy right-handed adults (age range of 19–33 years, 25 females) took part in the same DL experiment, this time with concurrent EEG recording from 61 electrodes. The study replicated the REA for SL pairs and the LEA for LS pairs based on percentage of correct reports in each ear. The analyses of cortical activity at around 100 ms after stimulus onset focussed on left temporal, central, and right temporal regions. Neither N1 amplitudes or latencies were in line with the systematic differences revealed by the behavioural data leading the authors to conclude that behavioural asymmetry could possibly reflect both stimulus-driven and strategy-driven factors. It is noted that the parameters of this electrophysiological investigation, like any other, were highly specified (and therefore perhaps somewhat limited).

While DL studies seldom report reaction times or overall error rates, Sandmann et al. (2007) found that, of the four VOT

[2] Note this refers to studies of VOT during dichotic listening. There is, of course, a long history of research on acoustic features of consonants such as VOT (e.g., Frye et al., 2007; Hutchison, Blumstein, & Myers, 2008; Liberman, 1957; Lisker & Abramson, 1967; Repp & Liberman, 1987, amongst many others).