Tone processing has been used as a vehicle for investigating numerous questions concerning the neurophysiology of basic auditory perception, language, and music. Behavioral, dichotic-listening, lesion, and functional and anatomical neuroimaging studies have shed light on various debates of the neural organization of perceptual information. From our early fascination of hemispheric specialization to more recently the neurogenetics of language, the study of tone has always provided key evidence for our most current understanding of the brain. Here, we provide a historical overview of the key issues and present findings from our laboratory from the past decade that contribute to these debates.

Index Terms: lexical tone, auditory perception, neurophysiology, neurogenetics

1. Introduction

More than four decades ago, Van Lancker and Fromkin [1] published their seminal study in the first volume of the *Journal of Phonetics* examining the neurophysiology of lexical tone. In the next three decades and possibly longer, the topic of hemispheric specialization of lexical tone that they explored continued to be a dominant theme in the neuroscience of language. In 2002, a review paper was published focusing on three decades of research on hemispheric specialization of lexical tone [2]. Advances in neuroimaging techniques in the past 10 to 15 years have made it possible for neurolinguists to explore a wider range of topics beyond hemispheric specialization.

Here, we highlight the major issues in the neurophysiology of lexical tone examined by research in the last four decades, including hemispheric specialization, stages of neural processing, and neural plasticity. Before addressing these issues, we give a brief overview of the auditory system as it pertains to lexical tone. We conclude this article by discussing new areas of research.

2. The Auditory System

Sound enters the auditory system through the external ear, or pinna, which funnels sound into the auditory canal. Sound waves vibrate the tympanic membrane and these vibrations are then transmitted to the oval window of the cochlea via the bones of the middle ear (malleus, incus and stapes). Mechanical energy (vibrations) is transduced into electrochemical neural potentials by hair cells in the organ of Corti, which resides atop the basilar membrane within the cochlea. Projections from the cochlea form the auditory nerve, whose fibers enter the brainstem. The auditory pathways ascend through a number of brainstem nuclei (see Fig. 1), including the cochlear nuclei in the medulla, and the superior olivary nuclei in the pons. Ascending projections then arrive at the inferior colliculus of the midbrain and the medial geniculate body of the thalamus and are then relayed to auditory cortex. In addition to these ascending projections, there are descending projections from auditory cortex via the corticofugal pathway back to the medial geniculate body and inferior colliculus, as well as other descending connections to...
the brainstem nuclei. Although not typically described as part of the auditory system, the inferior frontal gyrus (IFG) performs important language functions, and is active during linguistic decision making.

Sensitivity to increasing stimulus complexity is observed from the auditory periphery to the auditory cortex. While neurons towards the auditory periphery are sensitive to non-dynamic frequencies, neurons in the brainstem nuclei and in the inferior colliculus in particular are sensitive to frequency modulations (similar to lexical tones). Greater sensitivity to linguistic functions can be observed in the cortex.

3. Hemispheric Specialization

Hemispheric specialization of tone has been the predominant interest of neurolinguists seeking to understand the relationship between brain and tone for at least thirty years. That is, researchers seek to understand whether the left or right hemisphere of the brain subserves the production and/or perception of tones. The question of hemispheric specialization has long been a focus of cognitive neuroscience research, dating back Hippocrates’ time [3]. Centuries of research, using lesion, dichotic listening, and functional neuroimaging techniques have found left lateralization for language [4]. Relatively more recently, pitch processing has been attributed to the right hemisphere [5]. Lexical tones, being both pitch and language, became an interesting vehicle for studying brain organization. There are at least two hypotheses which research on the neurophysiology of tone could test. The functional hypothesis stresses the function of the acoustic stimuli, rather than the acoustic stimuli themselves. Thus, lexical tones, being linguistic in nature, would be lateralized to the left hemisphere. In contrast, the acoustic hypothesis focuses on the acoustic properties of the stimuli rather than their functions. Because lexical tones are frequency modulations, they are lateralized to the right hemisphere. It is worth emphasizing that it is lexical tones’ dual acoustic-functional properties that make testing these alternative hypotheses possible.

Van Lancker and Fromkin [1] were the first to publish a study examining hemispheric specialization of tones, as far as we are aware. A dichotic listening method was used. Because more nerve fibers from the ear eventually carry information to the contralateral side of the brain (see Fig. 1), via various intermediate centers, a right-ear advantage (REA) is taken as demonstration of left hemisphere specialization (and vice versa) in a perception task in which different stimuli are presented to the two ears. In their seminal study, Van Lancker and Fromkin presented Thai and English listeners with Thai words and hummed versions of the same words that differed in pitch. They found only an REA when Thai listeners listened to Thai words, suggesting a left hemisphere specialization of lexical tones. English listeners did not show an REA when they listened to Thai words. Neither groups showed a significant ear advantage when listened to hums. This study provides important first evidence for left hemispheric specialization for tones. It is important to note, however, that the Thai condition is the only one in which words were presented. Thus, the REA observed could be interpreted as a lexical (word processing) effect rather than a lexical tone effect per se. This potential lexical confound is important to note as more recent studies are reviewed.

Lesion studies are important sources when considering hemispheric specialization. A loss-of-function lesion on either side of the brain could be interpreted as hemispheric function on the injured side. Twenty production and 16 perception studies were reviewed in Wong [2] and the conclusion at the time was that the results were mixed at best (see Table 1 for a sample of these studies that investigated lexical tone). For example, the vast majority of the lesion studies have insufficient control. Often times, impaired functions could be due to a general rather than specific brain injury. Because several reviews have been published focusing on lesion studies [2], [6], [7], we will not discuss those studies at length here.

<table>
<thead>
<tr>
<th>Study (subjects)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production</strong></td>
<td></td>
</tr>
<tr>
<td>Packard [8] (8NF, 8N)</td>
<td>N &gt; NF</td>
</tr>
<tr>
<td>Gandour et al. [9] (9F, 6NF, 12R, 20N)</td>
<td>N = R = F = NF</td>
</tr>
<tr>
<td>Gandour et al. [10] (9F, 5NF, 13R, 20N)</td>
<td>N = R &gt; F &gt; NF</td>
</tr>
<tr>
<td><strong>Perception</strong></td>
<td></td>
</tr>
<tr>
<td>Eng et al. [12] (5L, 5N)</td>
<td>N &gt; L</td>
</tr>
<tr>
<td>Gandour &amp; Dardarananda [13] (3NF, 1F, 1R,1N)</td>
<td>R = N &gt; L</td>
</tr>
<tr>
<td>Hughes et al. [14] (12R, 7N)</td>
<td>N = R</td>
</tr>
<tr>
<td>Brådvik et al. [15] (20R, 18N)</td>
<td>N &gt; R</td>
</tr>
</tbody>
</table>

N = neurologically normal  
R = right hemisphere damaged  
L = left hemisphere damaged  
NF = nonfluent aphasic (L)  
F = fluent aphasic (L)

Results: ‘=’, equal performance; ‘>’, subjects could identify stimuli better than, or others identified subjects’ production better than.

Advances in neuroimaging technology have allowed for a new generation of research studies that have provided greater anatomical details about brain functions beyond lateralization. In terms of lexical tones, Gandour and colleagues pioneered a line of research in which functional neuroimaging was used to study lexical tones. In a first study, Gandour et al. [16] used Positron Emission Tomography (PET) and asked Thai and English listeners to discriminate the pitch patterns of Thai words and low-pass speech-filtered tone stimuli, and found activity in the left inferior frontal region only when the Thai subjects were listening to Thai words. This left-hemisphere result was taken as supportive of the functional hypothesis. However, as discussed above in the context of the Van Lancker and Fromkin [1] study, there is a potential confound of lexicality, for the Thai word condition was also a condition in which Thai listeners were listening to meaningful words.

In order to address the potential confound of lexicality, we argued that a condition in which discrimination of pitch patterns embedded in words from a non-tone language is needed. In Wong et al. [17] native English-speaking listeners and bilingual Mandarin-English listeners were asked to discriminate pitch patterns embedded in both Mandarin and English words. The pitch patterns resembled lexical tones in Mandarin. In this design, listeners were asked to listen to meaningful words across conditions and were asked to make judgment about pitch patterns, but the pitch patterns were only lexical for the Mandarin listeners when they were embedded
in Mandarin words. Pitch patterns embedded in English words were non-lexical and pitch patterns embedded in Mandarin words were non-lexical to native English-speaking listeners. Definitively supporting the functional hypothesis, activation in the left inferior frontal region was only observed when Mandarin listeners discriminated pitch patterns (lexical tones) in Mandarin words. In all other conditions, the activation was near the homologous area on the right.

4. Stages of Neural Processing

Research on the neurophysiology of tone in the past four decades has focused predominately on the cortex. As discussed earlier, the auditory pathway contains important structures before the cortex. For example, the inferior colliculus in the rostral brainstem contains neurons that are sensitive to frequency modulations. Because an increasing sensitivity to acoustic complexity is observed in the ascending auditory pathways, framing the question in regards to left and right brain is likely to be simplistic. The question should perhaps focus on what stimulus properties and functions are processed at which stage of the auditory pathway [18]. Such a focus would inevitably require examination of subcortical structures.

To date, subcortical studies of lexical tones have predominately focused on the inferior colliculus. The vast majority of these studies have used electroencephalography (EEG) to investigate how well pitch is encoded (to be discussed next), though recent studies have also begun to use more anatomically précised methods (e.g., fMRI) [19].

Because of its temporal resolution, EEG provides an invaluable means for observing stimulus encoding in the nervous system. A clear advantage to examining the encoding of stimulus pitch using EEG is that the frequency following response (FFR) is the brainstem evoked potential that reflects ensemble phase-locked responses to the stimulus [20], and encodes pitch with high fidelity [21] (see Figs. 2 and 3).

**Figure 2.** A sample electrode montage used in an FFR acquisition study. Note that the electrophysiological experiment required to capture the FFR requires only a simple setup. In the example above, only three scalp electrodes are used: the mid-frontal electrode (Cz) actively records brainstem activity, the mastoid electrode (M) serves as a reference, and the mid-forehead electrode (Fpz) is ground.

**Figure 3.** Top, acoustic stimulus token of Cantonese mid-rising tone produced in /j2/ syllable. Middle, FFR showing onset at 0 ms and phase-locked response for duration of 200 ms stimulus. Note that the brainstem neural activity closely matches the temporal waveform of the eliciting acoustic stimulus token. Bottom, brainstem pitch tracking trajectories (squares) elicited by the same stimulus. Solid black line represents the estimated stimulus $f_0$ contour.

It is now clear that the auditory brainstem is subject to tuning by long-term auditory experiences. Krishnan et al. [22] were the first to demonstrate speaking a tone language has an effect of brainstem encoding. In their study, Mandarin and English listeners listened to Mandarin tones while evoked responses from the rostral brainstem were recorded. Different measures of faithfulness in encoding were used, but all pointed to the fact that Mandarin listeners encoded Mandarin tones more faithfully than English listeners who did not speak Mandarin.

Following this first study, we asked whether or not long-term experiences with music would affect the encoding of Mandarin tones even if the listeners do not speak Mandarin [23]. Native English-speaking musicians and non-musicians who do not speak Mandarin listened to Mandarin tones while evoked responses from the brainstem were recorded. As with long-term linguistic experience, long-term musical experience...
also enhances the encoding of Mandarin tones even though the listeners did not speak Mandarin. Taken together, these two studies suggest that the encoding of lexical tone in the brainstem is modulated by long-term auditory experiences.

More about neural plasticity will be discussed below. Here, it is worth noting that studies now also suggest that the encoding of lexical tones in the brainstem can be changed even after short-term training [19], [24].

5. Neural Plasticity

For decades, questions of brain plasticity have concerned whether or not the brain can change, who can be changed, what can be changed, and the neurobiological mechanisms of change. In the past decades, the rise of translational research has pushed the scientific community to focus on the human brain. Language is a vehicle for examining changes in higher-level brain functions. Lexical tone is especially valuable to study because it represents an aspect of language and foundational knowledge of neural processing of pitch is available.

The first study to examine neural plasticity of lexical tone was conducted by Wang et al. [25]. Six Mandarin tone learners who spoke English as their native language completed a tone training study and neural responses were measured before and after training. Compared to pre-training, an expansion of activation especially in left-hemisphere language-related regions was observed post-training. Though a small-scale study, this was the first study to provide evidence for neural changes resulting from lexical tone learning, and reaffirm the role of left hemisphere regions for lexical tone even after shorter term language experience.

Lexical tone learning can also be used for examining newer questions in the era of translational research. For example, more recent studies attempt to answer how much are mechanisms of plasticity tied to behavioral performance/learning success, and whether or not neural (and neurogenetic) measures could “predict” learning success, even before training has commenced. In the past several years, we have conducted numerous studies on lexical tone learning aiming to address these newer questions. In many of the studies, we trained native English-speaking learners who had no previous exposure to lexical tones to use Mandarin tones to distinguish word meanings in the laboratory for multiple days. For example, in a picture-word matching paradigm, learners learned to use lexical tones to contrast word meaning (see Table 2 for examples of pseudowords varying in Mandarin tones).

Table 2. Words differing in Mandarin tones (1, 2 and 4) used in a learning experiment (e.g., Wong & Perrachione [26]).

<table>
<thead>
<tr>
<th>Tone 1 Meaning</th>
<th>Tone 2 Meaning</th>
<th>Tone 4 Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>pesh1 “glass”</td>
<td>pesh2 “pencil”</td>
<td>pesh4 “table”</td>
</tr>
<tr>
<td>dree1 “arm”</td>
<td>dree2 “phone”</td>
<td>dree4 “cow”</td>
</tr>
<tr>
<td>ner1 “boat”</td>
<td>ner2 “potato”</td>
<td>ner4 “dog”</td>
</tr>
<tr>
<td>vece1 “hat”</td>
<td>vece2 “tape”</td>
<td>vece4 “piano”</td>
</tr>
<tr>
<td>nuck1 “brush”</td>
<td>nuck2 “tissue”</td>
<td>nuck4 “bus”</td>
</tr>
<tr>
<td>fute1 “shoe”</td>
<td>fute2 “book”</td>
<td>fute4 “knife”</td>
</tr>
</tbody>
</table>

In one study, we found that the neuro-signature of successful learning is cortical streamlining; that is, while less successful learners activated a diffused brain network including the right frontal lobe, successful learners activated a focused area in the left temporal lobe to a larger extent [27]. We also found numerous predictive factors of learning success, which we measured before training. For examples, the volume of Heschl’s Gyrus [28], integrity of white matter tract projecting from the posterior temporal cortex ventrally [29], and potentially polymorphism of the ASPM gene [30].

6. Conclusions and Future Directions

With the groundwork set by Van Lancker and Fromkin [1], and many earlier studies examining the hemispheric specialization of tones, studies in the past decade have begun to explore complex questions about the neurophysiology of lexical tones. These more recent studies have asked whether or not subcortical structures are integral to how lexical tones are processed in our nervous system, which in turn have led us to consider where, when, and what acoustical and functional properties of lexical tones are processed. Recent studies have also taken a translational approach to try to use lexical tone as a vehicle to examine the neuro-signatures of training success as well as pre-training predictors of learning.

Many research topics await further investigations. In terms of the basic neural mechanisms, studies should carefully address how levels of neural structures and networks handle specific tonal processes, taking into consideration physical properties and functional contexts. There is a large gap in the literature on the neural development of tones: how does an infant’s nervous system develop to handle tonal information. Future translational research should consider what are the communicative consequences of tonal processing/production disorders and how they can be remediated.

7. Acknowledgements

We would like to thank the organizers of TAL 2014 for the invitation to present a keynote address. Our research is supported by the Liu Che Woo Institute of Innovative Medicine at The Chinese University of Hong Kong, the Global Parent Child Resource Centre Limited, the US National Institutes of Health grant R01DC008333, National Science Foundation grant BCS-1125144, the Research Grants Council of Hong Kong grant 477513, and the Health and Medical Research Fund of Hong Kong grant 01120616 awarded to P.W.

8. References


