

Nicol T and Kraus N (2005). How can the neural encoding and perception of speech be improved? In: *Plasticity and Signal Representation in the Auditory System*. Syka J and Merzenich MM (eds.), Kluwer Plenum, New York, pp 259-270.

## **How Can The Neural Encoding and Perception of Speech Be Improved?**

TRENT NICOL\*, and NINA KRAUS\*#

*\*Auditory Neuroscience Laboratory, Department of Communication Sciences, <http://www.communication.northwestern.edu/csd/research/brainvolts/> and #Departments of Neurobiology & Physiology and Otolaryngology, Northwestern University, Evanston, IL, USA*

### **1. INTRODUCTION**

A major focus of our Auditory Neuroscience Laboratory is gaining an understanding of the physiology underlying speech encoding. Imprecise encoding of speech sounds at the neural level contributes to communication problems in adverse listening conditions. Populations such as the hearing impaired, the elderly, and individuals with auditory-based learning disabilities suffer from poor speech perception even under favourable listening conditions. Unfamiliar sounds and sound combinations present perceptual difficulties for people struggling to learn a foreign language.

In order to improve speech encoding, and ultimately perception, there are two broad courses of action. The first is training. The auditory system has a remarkable ability to reorganize in response to training in order to cope with impoverished—or unfamiliar—signals. Second, steps can be taken to increase the clarity of the signal, shifting the burden of improving encoding outward.

This report, first, will review some work on our lab demonstrating neural plasticity resulting from training in normal, adult populations. Next, physiological changes seen in learning-disabled children following a commercial auditory training program will be examined. Finally, taking a

broader definition of plasticity, the neural alterations effected by modifications to speech sounds themselves will be addressed.

## 2. BACKGROUND

The impact of discrimination training on speech-evoked cortical responses has been examined in our laboratory from multiple angles. The plasticity of the N1/P2 complex—signalling an acoustic event, and the mismatch response—signalling a change between two acoustic events, has been investigated. Particular attention was paid to the time course of neural plasticity and to its generalization to similar acoustic stimuli.

The mismatch response is an attention-independent metric of the physiological detection of acoustic change. Thus, it is useful in the study of training-associated auditory discrimination abilities. If a previously dormant neural population newly reacts to a formerly undetectable difference, it can be inferred that the training regimen has had an effect even in the absence of overt behavioural improvement. Additionally, the N1/P2 complex has been demonstrated to reflect acoustic properties present in complex stimuli such as speech. In particular, voiceless consonant-vowel speech syllables such as /ta/ have been demonstrated to elicit discrete responses to the aspiration of the unvoiced consonant and to the onset of the voiced vowel. Thus, the N1/P2 complex lends itself to the study of training-related discrimination improvements in this type of stimulus. A review of several studies observing altered physiology after directed speech-sound training follows.

### 2.1 N1/P2 Response Plasticity

Two pre-voiced /ba/ variants, with voice-onset times of –20 and –10 msec, were perceived as /ba/ by ten monolingual English speaking young adults. Subjects were successfully trained, over the course of nine days, to identify the two syllables as /mba/ and /ba/, respectively. In order to determine whether the N1/P2 complex demonstrated changes corresponding to the newly learned temporal cue, cortical responses were acquired to the two /ba/ variants before and after the training regimen. The N1/P2 complex, in response to both stimuli, increased in amplitude following training. Moreover, there appeared to be a relationship between the magnitude of the amplitude increase and the extent of improvement demonstrated behaviourally (Tremblay *et al.*, 2001).

## 2.2 Mismatch Response Plasticity

A series of mismatch response plasticity investigations was undertaken. First, to establish whether attention-independent neural encoding of sound structure is mutable, thirteen adult subjects underwent speech-sound discrimination training (Kraus *et al.*, 1995). Prior to training, the subjects' discrimination of two similar /da/ stimuli was at chance. Following six one-hour training sessions, the subjects' discrimination significantly improved. Using the same two /da/s to evoke a physiological response, the mismatch response increased in both duration and amplitude in nearly all subjects following the training regimen—demonstrating evidence of cortical neural plasticity.

In a similar design, this time with /ba/ variants as stimuli, a more thorough investigation of the temporal relationship between behavioural improvement and physiological change was undertaken (Tremblay *et al.*, 1998). The mismatch response was measured following each training session in ten subjects. Evidence of neural plasticity—increases in mismatch response duration and magnitude, decrease in onset latency—preceded behavioural improvement in half the subjects. In the remainder, it was concurrent; in no case did the changed behaviour precede the physiological manifestation.

Next, generalization of learning to stimuli other than those used in training was investigated (Tremblay *et al.*, 1997). Nine adult English-speaking subjects were trained to discriminate between and label a pair of pre-voiced bilabial stop consonants—a relevant cue in some languages, but not English. Mismatch responses were collected to the pair before and after nine 20-minute labelling training sessions, over five days. Also assessed behaviourally and physiologically, but not trained, was the ability to discriminate a pair of pre-voiced alveolar stop consonants sharing the same voicing time distinction as the trained bilabial pair. Both behavioural and physiological gains were observed in the untrained alveolar pair as well as the trained bilabial pair.

## 2.3 Laterality of Physiological Plasticity

In addition to the training-related cortical physiology changes outlined above, physiological responses demonstrate topographical reorganization. With right-ear stimulation, cortical speech-evoked P1 and N1 responses demonstrated a larger post-training increase on the right side (Tremblay and Kraus, 2002). The mismatch response, however, demonstrated a larger magnitude increase on the left (Tremblay *et al.*, 1997).

Thus, several studies have demonstrated that the neural encoding of sound structure is mutable by training. In addition, these pre-consciously recorded changes in neural encoding are evident prior to their behavioural expression and the occurrence of training generalization also is manifested physiologically. Different patterns of laterality in various physiological responses are discernible. Thus, there is considerable evidence of plasticity in the neural encoding of sound structure in auditory cortex. Does plasticity occur subcortically?

With the establishment of this foundation, we have the basis to refocus our attention, outside of the realm of normal adults, on to the more pressing concern of the language-impaired child. The twin goals of improved reading and better perception and understanding of spoken speech are paramount in the treatment of children with language-based learning disabilities. The ability to monitor the brain's response as training improves these skills will enable a better understanding of the mechanisms involved in language acquisition. In addition, because a chief component of auditory training programs is the selective enhancement of speech signals—for example lengthening of consonant-vowel formant transitions, directed studies examining the impact of such enhancements on the neural encoding of speech can be undertaken. In combination, these approaches enable us to assess both the efficacy of the program as a whole and the cue-enhancement technique in isolation.

### **3. COMMERCIAL AUDITORY TRAINING**

A comprehensive project studying physiological response differences between normal children and children with a variety of auditory-based learning disabilities has been ongoing in our laboratory. Differences in neural responses to speech syllables have been established between children with learning problems and normal controls, and relationships between behavioural speech-sound perception abilities and auditory physiology have been demonstrated. Having established norms for school-aged children on a variety of speech-evoked cortical and subcortical responses, we have a good metric against which to evaluate the mutability arising from auditory training in children with learning problems (Kraus *et al.*, 1996; Kraus *et al.*, 1999).

A unique opportunity to investigate neural plasticity to speech sounds in such children has arisen due to the recent popularity of commercial computer-based auditory training programs (Diehl, 1999; Merzenich *et al.*, 1996; Morrison, 1998; Tallal *et al.*, 1996). Children undergoing this type of training comprise a heterogeneous population. This provides an opportunity to study neural plasticity among individuals with a wide variety of learning

problems, and of practical importance, to examine physiological patterns in children for whom training may result in a variety of outcomes, thus leading to a potential objective metric presaging probable success.

Twenty-seven subjects with auditory-based learning problems were tested on a number of behavioural and physiological measures prior to and following independently administered *Earobics* (*Cognitive Concepts, Inc., Evanston, IL, USA*) training. *Earobics* consists of directed exercises that incorporate phoneme discrimination, auditory memory, sequencing, attention, rhyming and sound blending skills aimed at improving phonological awareness and ultimately reading. The behavioural and physiological test battery was administered to monitor changes in and relationships among these tests resulting from training. The test battery also was administered twice to fifteen non-trained controls spanning a similar time interval.

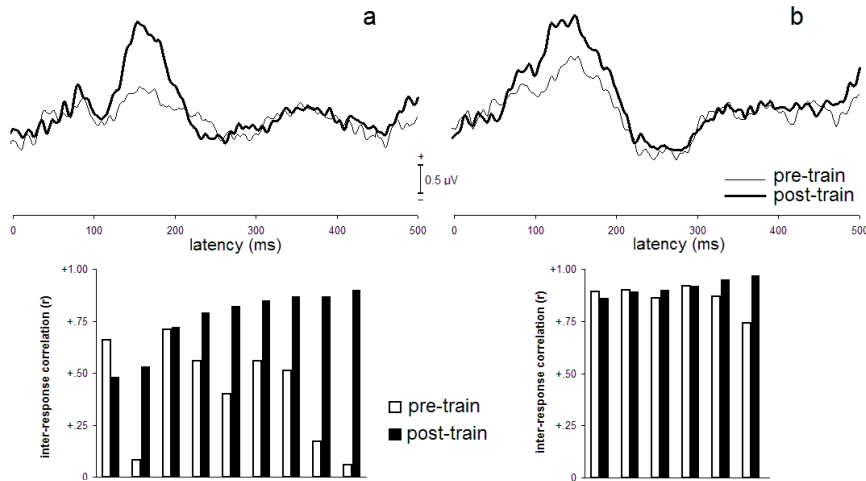
The behavioural battery included subtests of the Woodcock-Johnson Psycho-Educational Battery (Woodcock and Johnson, 1977; Woodcock and Johnson, 1989) and the Wide Range Achievement Test (Wilkinson, 1993). Listening tasks included sentence perception in noise and just noticeable differences for various consonant-vowel syllable continua in quiet and noise. The physiological battery consisted of subcortical and cortical response recording to /da/, both in quiet and noise, and mismatch responses to a near-threshold /da-ga/ pair in quiet.

Following training, in addition to improving on a number of behavioural tests associated with speech discrimination, comprehension and phonetic awareness, the experimental group demonstrated some physiological changes not seen in the controls. Responses of both cortical and subcortical origin were altered by training.

Cortical P2/N2 responses took on a more mature-looking pattern in quiet following training. Children with more mature cortical responses were most likely to improve on portions of the behavioural test battery. Furthermore, in noise, the amount of response degradation in comparison to the response in quiet, measured by inter-response correlation, was diminished (Hayes *et al.*, 2003). Figure 1 illustrates cortical responses in noise, recorded before and after training (top). Subjects whose inter-response correlations were low prior to training (left) demonstrated sharp increases in correlation values (bottom). Subjects with high inter-response correlations (right) prior to training maintained their strong response in noise following training. The cortical mismatch response was altered as well. Response topography shifted to a more left-dominant pattern following training.

The subcortical response to /da/ consists of a transient response to stimulus onset followed by a sustained frequency-following response to the harmonic aspects of the vowel (King *et al.*, 2002). Training affected the two

subcortical response components differently. The transient onset was unaffected by training; both latency and amplitude of this response were stable. However, the subjects whose onset response latencies were most delayed irrespective of test session demonstrated the most post-training improvement on cortical physiology and behavioural measures of auditory perception—potentially revealing an important screening tool. The sustained response, unlike the onset, did demonstrate evidence of modification following training. Quiet-to-noise inter-response correlations increased. Viewed in conjunction with the aforementioned improved cortical inter-response correlations, a consequence of training appears to be improved neural timing in noise in both the auditory brainstem and cortex.



*Figure 1.* Grand average cortical response to /da/ in noise and individual inter-response correlations. Subjects with poor pre-train inter-response correlations (a) showed a marked morphological change (top) in their response that accompanied their increased correlations (bottom). Subjects whose inter-response correlations were high prior to training (b) demonstrated little change in response morphology and correlations.

Thus, commercial auditory training appears to have a measurable effect on speech-evoked physiology. In particular, the inter-response correlations, both in subcortical and cortical physiology, indicate a movement toward greater precision in neural population timing. Moreover, the source of the improvement appears to lie in enhanced responses to speech signals that are masked with noise.

#### 4. STIMULUS CUE-ENHANCEMENT

In school-age children, certain cortical speech-evoked responses and speech perception tasks are especially affected by the introduction of background noise to the stimuli (Cunningham *et al.*, 2001; Wible *et al.*, 2002). For some learning-disabled children, naturally produced clear speech enhancements are sufficient to improve sentence perception in noise performance equal to normal controls (Bradlow *et al.*, 2003), and speech enhancement techniques are a critical component of auditory training programs (Merzenich *et al.*, 1996). Thus, a line of research has been pursued that investigates physiological response modifications arising from manipulation of the speech signal itself rather than response modifications arising from remediation.

In order to investigate enhanced speech on a syllable level, two cue-enhancement strategies were applied to a 40-step /ada-aga/ stimulus continuum that was characterized as “conversational.” First, the stop-gap duration was increased by 80 ms, and second, the amplitude of the consonant burst was increased by 10 dB. A third continuum variant combined both enhancement strategies. The vowel following the stop was not modified. These modifications are common in naturally produced “clear” speech (Picheny *et al.*, 1986).

In nine children with auditory-based learning problems (LP) and nine normal-learning (NL) controls, just noticeable differences were established on the conversational continuum presented in both quiet and continuous background noise. The groups were equivalent in their ability to discriminate among members of the continuum in quiet, but the LP children were significantly poorer at the task than the controls when background noise (+5 dB signal-to-noise ratio) was added. Both cue enhancement strategies afforded behavioural improvement in noise. In particular, the continuum utilizing enhanced burst amplitude restored the LP subjects’ performance to a level equal to the controls (figure 2a).

A similar pattern was seen in evoked response patterns<sup>1</sup>. In quiet, the cortical P2/N2 response complex to the conversational stimulus was equivalent in amplitude between the two groups. However, the addition of background noise, while diminishing the response amplitude in both groups, suppressed the response to a greater degree in the LP children. Likewise, subcortical responses that were alike between groups in quiet were

<sup>1</sup> Responses were recorded to the endpoint /da/ rather than /ada/. In assessing the behavioural response enhancements effected by the two cue enhancement strategies, it was found that enhancing the consonant was more effective than elongating the stop gap. Thus, it was decided to concentrate on the former for subsequent physiology recordings, enabling the removal of the leading vowel.

differentially affected upon the addition of background noise to the conversational stimulus. In noise, both transient (later in LPs), and sustained (reduced high-frequency phase locking; poorer stimulus-to-response correlations in LPs) components of the response were significantly different from the controls.

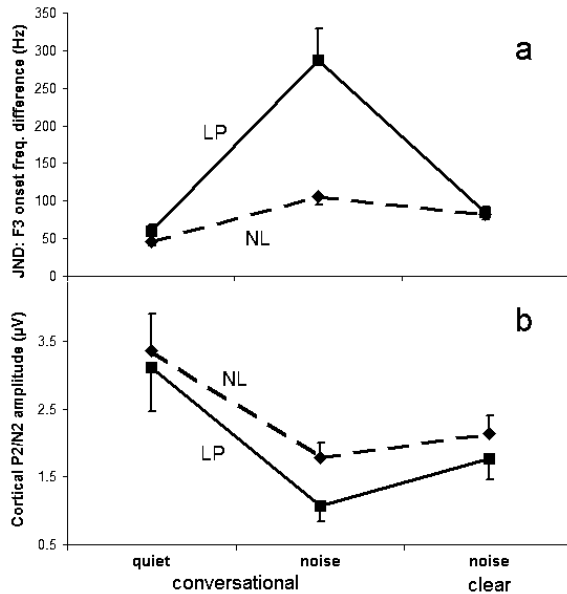


Figure 2. Behavioural discrimination scores (a) and cortical physiology (b). With conversational speech, LP subjects are more affected than normal controls by noise in both behaviour and physiology. Cue-enhanced, clear speech returns their behaviour and physiology to normal. Adapted from Cunningham *et al.* (2001).

Physiological responses to the cue-enhanced stimulus in noise were more robust for the LP subjects. Their cortical P2/N2 response amplitudes were no longer significantly different from the controls' (figure 2b). The subcortical onset response resumed a normal latency. However, the later, sustained, component of the subcortical response did not rebound with cue enhancement. That the sustained frequency-following response was unchanged by the stimulus cue-enhancement is unsurprising. Because the vowel was unaltered acoustically by cue-enhancement, it served as a useful control; helpful in confirming that the changes seen in responses to the sound onset were due to cue-enhancement alone, and were not attributable to changes in subject state or earphone placement over time.

An accompanying study (Cunningham *et al.*, 2002), using a guinea pig model to investigate near-field response patterns in the inferior colliculus, medial geniculate nucleus and auditory cortex, demonstrated that both the



disruption of background noise the benefit afforded by cue-enhancement was greatest at the cortex.

The behavioural improvement offered by stimulus cue-enhancement appears to have a direct link to more accurate neural representation of acoustic events—both cortically and subcortically.

Recent technological advances have permitted the elongation and amplification of certain speech cues to occur almost instantaneously. Consumer products are in development that process radio broadcasts into cue-enhanced speech in real time.

## **5. CONCLUSION**

Accurate speech-sound encoding and perception depends on precise timing of neural events. A deficit in speech-sound perception may be cognitive in origin, but in many instances, imprecise afferent encoding is its source. It has been demonstrated that physiological responses signalling the proper encoding of speech are measurable, and deficits have been discerned in both the auditory cortex and in subcortical regions.

Importantly, this preconscious encoding of sound is plastic. Auditory training has long been used to improve perception. With physiological recording, it is possible to track where neural reorganization has occurred, providing us with insight into the nature of auditory system plasticity as well as an unbiased gauge of training success.

Training and stimulus cue enhancement's impact on optimising neural timing to acoustic sound structure leads to speculation about their impact on learning other types of sounds. The malleability of encoding of sound within the auditory pathway suggests approaches that could be applied more generally in other instances where improved perception of sound is desirable, such as learning music or foreign languages.

## **ACKNOWLEDGEMENTS**

We are grateful for the scientific contributions of Erin Hayes, Bradley Wible, Jenna Cunningham, Kelly Tremblay, Cathrine Warrier and Cynthia King.

This research was supported by NIH NIDCD R01-01510.

## REFERENCES

- Bradlow, A. R., Kraus, N., and Hayes, E., 2003, Speaking clearly for children with learning disabilities: Sentence perception in noise. *J. Speech Lang. Hear. Res.* **46**: 80-97.
- Cunningham, J., Nicol, T., King, C. D., Zecker, S. G., and Kraus, N., 2002, Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hear. Res.* **169**: 97-111.
- Cunningham, J., Nicol, T., Zecker, S. G., and Kraus, N., 2001, Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clin. Neurophysiol.* **112**: 758-767.
- Diehl, S., 1999, Listen & Learn? A software approach review of Earobics. *Language, Speech and Hearing Services in Schools* **30**: 108-116.
- Hayes, E., Warrier, C. M., Nicol, T., Zecker, S. G., and Kraus, N., 2003, Neural plasticity following auditory training in children with learning problems. *Clin. Neurophysiol.* **114**: 673-684.
- King, C., Warrier, C. M., Hayes, E., and Kraus, N., 2002, Deficits in auditory brainstem encoding of speech sounds in children with learning problems. *Neurosci. Lett.* **319**: 111-115.
- Kraus, N., Koch, D. B., McGee, T. J., Nicol, T. G., and Cunningham, J., 1999, Speech-sound discrimination in school-age children: psychophysical and neurophysiologic measures. *J. Speech Lang. Hear. Res.* **42**: 1042-1060.
- Kraus, N., McGee, T., Carrell, T., King, C., and Tremblay, K., 1995, Central auditory system plasticity associated with speech discrimination training. *J. Cogn. Neurosci.* **7**: 27-34.
- Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., and Koch, D. B., 1996, Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* **273**: 971-973.
- Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., and Tallal, P., 1996, Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* **271**: 77-81.
- Morrison, S., 1998, Computer applications: Earobics Pro. *Child Language Teaching and Therapy* **14**: 279-284.
- Picheny, M. A., Durlach, N. I., and Braidia, L. D., 1986, Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *J. Speech Hear. Res.* **29**: 434-446.
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., and Merzenich, M. M., 1996, Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* **271**: 81-84.
- Tremblay, K., Kraus, N., Carrell, T. D., and McGee, T., 1997, Central auditory system plasticity: generalization to novel stimuli following listening training. *J. Acoust. Soc. Amer.* **102**: 3762-3773.
- Tremblay, K., Kraus, N., and McGee, T., 1998, The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport* **9**: 3557-60.
- Tremblay, K., Kraus, N., McGee, T. J., Ponton, C. W., and Otis, B., 2001, Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear.* **22**: 79-90.
- Tremblay, K. L., Kraus, N., 2002, Auditory training induces asymmetrical changes in cortical neural activity. *J. Speech Lang. Hear. Res.* **45**: 564-572.

- Wible, B., Nicol, T. G., and Kraus, N., 2002, Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. *Clin. Neurophysiol.* **113**: 485-494.
- Wilkinson, G., 1993, *Wide Range Achievement Test-3*. Jastak Assoc., Wilmington, DE.
- Woodcock, R., Johnson, M., 1977, *Woodcock-Johnson Psycho-Educational Battery. Tests of Cognitive Ability*. DLM Teaching Resources, Allen, TX.
- Woodcock, R., Johnson, M., 1989, *Woodcock-Johnson Psycho-Educational Battery-Revised . Tests of Cognitive Ability*. DLM Teaching Resources, Allen, TX.