

# When the Vocal Chords Vibrate but No One Listens: German Perception of English Word-Final Obstruents

Sarah Knee and Ruben van de Vijver  
University of Potsdam

## 1. Introduction

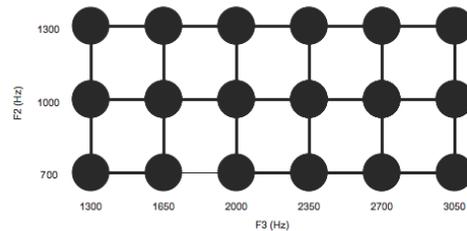
In the same way that we speak with an accent, we *listen* with an accent. Each of our perceptual systems is specialized for our native language. This perceptual specialization occurs during first language acquisition (e.g., Trehub, 1976, Jusczyk et al., 1994, Polka & Werker, 1994, Werker & Tees, 2002). It is clear that children are able to attend to the patterns and statistical distributions of their ambient language (Saffran et al., 1996, Maye et al., 2002). What is unknown is what is done with this information – is it used to restructure perception, or add additional processing layers to perception? This question has divided speech perception models into two broad categories: structure-changing and structure-adding. This paper examines these two theories, looking to the assumption, predictions, and experimental support they have received. This is followed by an experiment on German perception of English word-final voicing to fill a gap in the literature.

## 2. Comparing the Models: Assumptions, Predictions, and Support

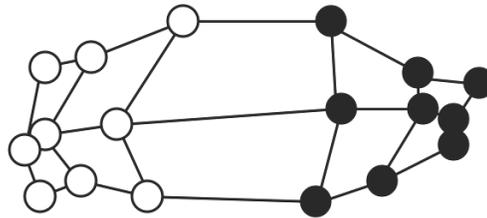
The majority of structure-changing models are phonetic. Initial speech perception is thought to be indistinct from general auditory perception. Exposure to an ambient language is assumed to alter perceptual space, either by the formation of categories or prototypes. Categories/prototypes are formed where children recognize meaningful contrasts in the linguistic input. Perception between these categories is enhanced, while perception within a category is diminished. Speech perception is then a single automatic, language-specific mapping from raw acoustic data to linguistic units, which is a natural consequence of the properties of the auditory system. An example of this is Kuhl's (2000) Native Learning Magnet (NLM). In this model, language perception and (mammalian) auditory perception begin the same (c.f. Kuhl & Miller, 1975, Hauser et al., 2001, Newport et al., 2004). Perceptual specialization begins as children attend to patterns and statistical distribution of their target language. Prototypical exemplars of the target language segments begin to form. These prototypes attract sounds within the category of the prototype, and repel other sounds (Grieser & Kuhl, 1989, Kuhl, 1991, Kuhl, 2000). This results in a bent perceptual space, as illustrated in (1) and (2) (modified from Iverson et al., 2003). The first figure illustrates an unaltered perceptual space; the second, an English perceptual space. White circles represent exemplars of the /r/ category, and black the /l/ categories. The perceptual space between like exemplars is shrunk while the space between

unlike exemplars is stretched.

(1) F2 and F3 formant frequencies of unaltered perceptual space



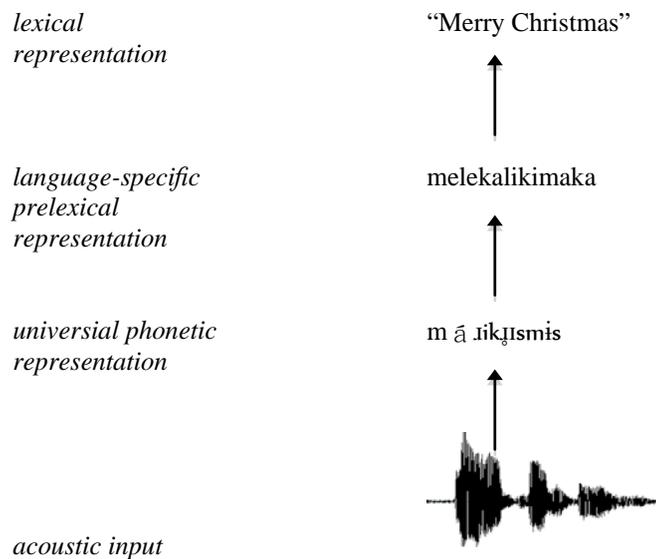
(2) F2 and F3 formant frequencies of English L1 perceptual space



Structure-adding models are often phonological. Initial speech perception is assumed to arise from a universal, speech-specific level of perception. Perception at this level is able to distinguish between any linguistic contrast in any natural human language. Language-specificity arises not from reorganization, but from added structural levels of perception namely, a prelexical filter and lexicon. Perhaps the most well known example is the filter model described by Dupoux and Peperkamp (2002). Initial speech perception occurs at the universal level, where linguistically unimportant information (speech rate, sex of the speaker, background noise, etc.) is reduced. Linguistically important information is then processed by the language-specific filter. The filter contains only first language (L1) phonemes assembled according to L1 phonotactics and so filters out nonmeaningful/redundant L1 contrasts before they reach the lexical level. The remaining information is now available for the lexical processing level. This is illustrated in (3).

The figure below is represents Hawaiian perception of English “Merry Christmas”. After the speech signal is received, it goes through the universal phonetic prelexical representation where linguistically irrelevant information is diminished. Then this information is further processed by a Hawaiian language-specific filter. In this case, vowels are added between consonant clusters and word-finally to accommodate Hawaiian phonotactics and non-native phonemes are matched to the closest native phonemes. This information is finally processed at the lexical level, where meaning is associated with the words.

- (3) Hawaiian perception of English “Merry Christmas”; Dupoux and Peperkamp (2002) Filter model



The assumptions and predictions of these model types stand in stark opposition to each other. Structure-changing models assume that speech perception results from general auditory processing (i.e., is not speech-specific); structure-adding models claim it is specific to speech. Structure-changing models assume that initial speech processing changes fundamentally during first language acquisition; structure-adding models insist that universal speech perception remains unchanged. While not explicitly stated, the lack of mention of lexical influence suggests structure-changing models assume if a distinction can be perceived, it can be used in (all) further stages of speech perception; structure-adding models hold that this distinction can be later filtered out. Indeed, the only aspect on which the models agree is that speech perception becomes language-specific during first language acquisition. The same dichotomy exists between the predictions of the two model types. Structure-changing models predict a difference in initial perception from one language group to another; structure-adding models predict this will not happen. Structure-adding models further predict a change in initial perception and perception at the lexical level. Such an occurrence cannot be supported by structure-changing models.

Despite the opposition between assumptions and predictions, both model types have been substantiated in the literature. Supporting structure-changing models is the research by Dehaene-Lambertz (1997), Iverson and colleagues (2003) and Best and colleagues (2001). In an ERP study, Dehaene-Lambertz (1997) found that only native place of articulation contrasts elicited a strong mismatch negativity (MMN). Participants were presented with blocks of four syllables. The first three were always identical, while the last was (1) physically

identical, (2) physically different but within the same native category, or (3) physically different and part of a different native category. Only the last case evoked an MMN response. Both Iverson and colleagues (2003) and Best and colleagues (2001) demonstrated that listeners from different language groups perceived contrasts differently using behavioural studies. Iverson et al. compared the perception of synthesized English /l/ and /r/ (differing only in the second and third formant (F2 and F3) values) by monolingual Japanese, German, and English native speakers. Japanese listeners used information pertinent to Japanese processing (namely F2), German to German (F2 and F3), and English to English (F2 and F3). Using this information, the authors were able to construct mental perceptual maps, of which, the English is presented earlier in this paper. Best and colleagues (2001) demonstrated that native English speakers' perception of Zulu and Tigrinya depended on English contrasts. Where the listener was able to assimilate a particular sound to an English category, perception was good. Where a sound was a poor exemplar of an English category, perception was poor.

Phonological models have received support both of a universal level of processing and an intermediate filter. Rivera-Gaxiola and colleagues (2000) conducted a different place of articulation ERP study using a passive listening mismatch paradigm. English listeners heard four kinds of syllables, two labial [ba] syllables (within category contrasts), dental syllables ([d̪a]), and retroflex syllables ([ɖa]). The ERP results revealed that the participants were able to perceive all of these contrasts, supporting a universal level of perception. Filter models have also been supported by appealing to perceptual illusions. Dupoux and colleagues (2001) explained mirage effects in Japanese consonant cluster perception by using a filter model. In two experimental tasks, transcription and lexical decision, the authors demonstrated that native listeners of Japanese, a language without consonant clusters, perceived an [u] between consonant clusters. During the transcription task, participants transcribed all consonant clusters with an intervening [u], regardless of whether or not this made a lexical word. For example, *sokdo* was transcribed as the Japanese real word *sokudo*, while *mikdo* was transcribed as *mikudo* despite the existence of the real word *mikado*. During the lexical decision task, both stimuli like *sokdo* and *sokudo* were identified as real words (at the same rate) and words like *mikdo* and *mikudo* were identified as non-words. This clearly demonstrated a prelexical as opposed to lexical perceptual effect. A Japanese prelexical filter contains only native phonemes assembled according to Japanese phonotactics. A word like *sokdo* is provided with an epenthetic [u] in order to fit Japanese phonotactics. Filter models were also used by Dupoux and colleagues (1997, 2008) and Peperkamp and colleagues (1999) to explain phonological deafness of suprasegmentals. They compared stress perception by French late-learners of Spanish with Spanish and French monolinguals. Using a prelexical discrimination task, they revealed that all of these listeners were able to perceive contrastive stress, even though this only existed in Spanish, thus supporting a universal processing level. Interestingly, only the Spanish listeners were able to use this perception in a lexical task; both groups of French listeners performed at chance level. This change from initial to lexical perception fits the prediction

made by positing a filter.

Several things stand out from these studies. First, structure-changing models have all been supported by tasks tapping *prelexical* perception of *segmentals*. While structure-adding models have garnered support from studies on both segmental and suprasegmental perception, only the latter has been tested at the lexical level. Moreover, the ERP experiment testing segmental perception is similar to the Dehaene-Lambertz (1997) ERP experiment, but yielded contrary results. It is beyond the scope of this paper to account for why this might be so. Rather, this experiment was included to make the point that while there has been research on segmental perception supporting structure-adding models, they don't involve a lexical task, and furthermore remain inconclusive. This leaves us with a gap in the literature: there has yet to be a study on prelexical and lexical perception of a segmental contrast. Without filling this gap, one is not able to say if the difference in results is due to a methodological issue, or simply a processing difference between segmentals and suprasegmentals. Indeed, neuroimaging studies have shown that foreign suprasegmentals tend to activate the right or both hemisphere(s), whereas foreign segmentals activate the left (Gandour et al., 1998, Klein et al., 2001). Perhaps this processing distinction carries over to lexical discrimination. It is possible that while suprasegmentals are "filtered" after the initial stage of prelexical perception, no such filter exists for segmentals. This leaves us with our two research questions:

1. Can there be a perceptual change of segmentals from prelexical to lexical levels of processing?
2. If so, does this change result in phonological deafness similar to the documented stress deafness?

If the first question is answered negatively, one would be left without cause to posit such an intermediate level, at least for segmentals, lending support to phonetic models. A yes to the first question would confirm an intermediate level of processing between initial and lexical processing and thus support phonological models. A further affirmative answer to the second research question would lend support to Dupoux and Peperkamp's filter model.

This is where our current research comes in. Dupoux and Peperkamp (2002) explain that segmental as well as suprasegmental deafness is predicted under the filter model, providing German perception of word-final voicing as a potential topic for further research. German exhibits final de-voicing – all syllable-final obstruents are produced without voicing. This makes syllable/word-final voicing information redundant, which a German-learning infant, attending to the patterns in his/her ambient language, would realize. Because syllable/word-final voicing is redundant, it should not be encoded in the German-specific filter and, therefore, should also not be part of the lexical information available to German natives. This should carry forward into German perception and storage of foreign words. They should not be able to use (or even perceive) syllable/word-final voicing information for lexical tasks. A structure-changing model would make the opposite prediction. Since German does contain a voicing contrast among obstruents (word-initially), German natives should have prototypes/categories built that will enhance this

perception. They should also be able to use this perception for foreign words. To test these predictions, we examine the German perception of English word-final voicing using both a prelexical and lexical task.

### 3. Current Research

German English listeners are ideal for testing segmental perception due to German and English phonologies with regard to obstruent voicing and vowel length. Both German and English contain similar voicing contrasts between obstruents word-initially; voiced obstruents are produced with short-lag aspiration, and voiceless obstruents with long-lag aspiration (Iverson & Salmons, 1995). Word-finally, however, this contrast is lost in German through final devoicing (FD). All word-final obstruents are produced voiceless. English, on the other hand, retains a contrast between voiced and voiceless obstruents. This contrast *can* be realized by a distinction in degree of aspiration of word-final obstruents (with voiceless consonants having greater aspiration), but the main cue is the length of the preceding vowel. Vowels are lengthened before syllable-final voiced consonants (Raphael, 1972). Obstruent voicing in German and English is summarized in the following chart:

#### (4) Realization of Voicing Contrasts in German and English

Word-initial		
Language	Voiced	Voiceless
German	Short-lag Aspiration	Long-lag aspiration
English	Short-lag Aspiration	Long-lag aspiration

Word-final		
Language	Voiced	Voiceless
German	Long-lag aspiration	Long-lag aspiration
English	Long preceding vowel	Short preceding vowel (optional aspiration)

German also contains contrastive vowel length, although this contrast is used to distinguish among vowels rather than obstruents (Wiese, 2000). Having contrastive vowel length in the listener's native language has been shown to predispose the listener to perceive English voicing contrasts (Crowther & Mann, 1992). Under phonetic models, German listeners should be sensitive to both aspiration and length cues. Indeed, their perception should be fundamentally altered to enhance the perception of these contrasts. Hence, they should be enabled to efficiently perceive English obstruent voicing, even word-finally, and perception should not change from the prelexical to the lexical level. Phonological models predict that German listeners should have no difficulty perceiving any contrast during initial perception because this is subject to universal perception – if it can be perceived in one language, it can be perceived in all. Dupoux and Peperkamp's model predicts that because German voicing is redundant word-finally, voicing information should be filtered-out between the

prelexical and lexical level. This should reveal itself in the way of segmental deafness. We conducted two experiments to test these predictions, an ABX discrimination task to determine if German native listeners can perceive English word-final voicing and a speeded lexical decision task to see if they can use this perception for lexical decision.

#### 4. The Experiments

##### 4.1 Participants

Participants were 7 German learners of English. The age of the participants ranged from 19-35, with an average age of 24. One of the participants was employed in Berlin; the rest were students in Berlin or Potsdam. All had studied English in school starting between the ages of 10 and 12. They had at least 8 years of experience with English, and 1 had lived abroad in an English-speaking country for 12 months. There was also a control group of 5 native Canadian English speakers, ranging in age from 23-25, with an average age of 24.

##### 4.2 Experiment 1

The first experiment was designed to test if participants were able to perceive the English word-final voicing contrast. To do this, we employed an ABX discrimination task involving a voicing contrast.

###### 4.2.1 Materials

There were 160 trials containing three words, A, B, and X. There were four trial types, listed below. In 50% of the trials, A was a real-word and B a corresponding non-word. In half of these trials, X=A. In the other 50% B was a real-word and A a corresponding non-word. Again, X=A for half of these trials. X was never acoustically identical to A or B. Noise was added to the original item to make the item sound as if it was spoken quickly on a subway. The items were synthesized using the filter(one formant) option in Praat, Frequency 1000Hz and Bandwidth 100Hz and speeded to 80% of the original speed using Praat's lengthen option. This was done to ensure that participants could not rely solely on acoustic cues to distinguish the items but rather had to process the items phonetically.

###### (5) Trial types

A	x	B	x	X
real word	x	non-word	x	X=A
real word	x	non-word	x	X=B
non-word	x	real word	x	X=A
non-word	x	real word	x	X=B

There were 80 non-filtered test items (40 real English words and 40 corresponding non-words), 80 fillers (40 real English words and 40 non-words),

and 80 filtered items. All words were monosyllabic (of the form C(C)VC). There were 10 real-words for each relevant voicing contrast, p/b, t/d, k/g, and f/v, 5 of these ended with the voiceless consonant, and 5 with the voiced. The non-words differed from the real-words in the voicing of the final consonant. All words were recorded in a sound-attenuating booth by a native speaker of Canadian English. Originally we wanted to test all English stops and fricatives. The fricative pairs /ʒ/ and θ/ð could not be tested on account of the low frequency of the voiced counterparts. Thus we decided to test p/b, t/d, k/g, f/v, and s/z. We ran a pre-test with 5 native Canadian English speakers using 10 real words and 10 corresponding non-words for each phoneme, save /g/, for which there were only 8 of each, for a total of 196 items. In the pre-test, participants were asked to decide what they thought the last sound of each word was. They were trained prior to the test to provide the last *sound* and not the last *letter* of each word. They could take as much time as they wanted, and they were able to replay each word three times. Each item was repeated in the test three times, for a total of 588 trials. Any item that received three errors, either from a single speaker or across speakers, was removed from the materials. Unfortunately, this led to removing the s/z items as participants could not reliably decide if the final sound was a [s] or a [z]. The vowels of the remaining items were then measured to ensure that this was indeed a reliable cue for syllable-final obstruent voicing.

From the remaining items, the five items with the highest frequencies (as measured by hit counts on google search) were selected for each phoneme. All items received a hit count above 50 million except *stab* (22. 4 million), *gig* (40. 6 million), *shave* (13. 6 million), and *stove* (20. 5 million). These items were included in order to reach 5 items per phoneme. An informal preliminary vocabulary test with some beginner to intermediate learners of English suggested these would not pose a problem.

#### 4.2.2 Procedure

Participants were tested individually, with audio being presented over stereo headphones (Sony, MDR-V600). The experiment consisted of 80 trials, 10 practice, and 70 experimental. All trials were randomized in the experimental phase for each participant. Each trial consisted of three items, A, B, and X, presented 700ms after the other. Participants were asked to listen to each of the words, and decide whether the last word was more similar to the first or to the second word by using the touch pad to select the respective '1st' or '2nd' button which appeared on screen. No time limit was imposed. The next trial began 1000ms after they finished the first item. The experiment began with 10 practice trials in which the participants could familiarize themselves with the test. After this phase, participants were given time to ask questions if they had any. Otherwise, they clicked the screen to continue with the experiment phase. Only fillers were used in the practice phase. The experiment lasted about 20 minutes.

#### 4.2.3 Results and Discussion

On the surface the results appear obvious: both German and native English

listeners were able to successfully complete the ABX task, having error rates of 15% and 1% respectively. The breakdown of errors is shown in (6).

(6) Average Error Rates (%) on ABX Task

<b>Group</b>	<b>X=A</b>	<b>X=B</b>	<b>Total</b>
German	4	26	15
English	1	1	1

However, two additional observations stand out, namely that the English controls seem to have an over-all easier time with the task, and that the German English listeners appear to have an especially hard time with the conditions in which X=B. To test if these differences were significant, a subjects and items anova was conducted. For the subject analysis, the within-subjects factors were response type (whether X=A or X=B) and the between-subjects factor was language. For the items analysis the factors were reversed. The anova confirms that the variance between language groups is significant ( $F(1,37) p < .05$ ,  $F(1,14) p < .05$ ) as well as the response type ( $F(1,24) p < .05$ ,  $F(1,7) p < .05$ ). There was also an interaction between language and the response type ( $F(1,25) p < .05$ ,  $F(1,7) p < .05$ ).

The finding that there are significant effects of response type and language group are both quite interesting. Unfortunately there is nothing much to be said about the response type effect. Due to the test design one cannot say if the effect is due to positional issues (namely, to closeness of X to the correct response) or the items themselves. A similar experiment in which two item lists were made, where X(1-20)=A in the first list but B in the second, would be able to shed light on this issue. Since this was not done, nothing further will be said about it here. Fortunately, more consideration can be given to the language group effect. This experiment was designed to tap into the initial phonetic processing of speech input. Under the model of Dupoux and colleagues this would be the level of universal phonetic prelexical representation. At this level one would not expect a significant difference between language groups, yet this is indeed what is found here. Phonetic models are capable of explaining this difference by attributing it the language-specific processing of speech perception which exists at even the most basic stages. Although, this would beg the question as to what specific language differences attributed variance in performance. Both German and English listeners are sensitive to vowel length, as established in an earlier section and demonstrated in this experiment. Perhaps the differences could be attributed to the vowels themselves. Or perhaps the variation is not due to linguistic factors at all. After each experiment most of the participants commented on how they thought their performance was. Most of the German participants said that they felt the experiment was “too easy”. While they did not know what the experiment was designed to test, they all understood that their perception of a foreign language was being examined. With this knowledge they assumed that the experiment would be difficult for them and when they found that it was not, they thought there must be a trick to it. The English listeners of course did not have these thoughts going into the

experiment. It is conceivable that this non-linguistic difference contributed to the variance in performance. Perhaps a similar experiment with feedback during the practice phase could reduce or eliminate this effect. Still, the main question this experiment was conducted to answer is whether or not German English listeners are able to distinguish between English real and non-words differing only in word-final voicing. The answer to which is a resounding yes. Further discussion on these points will be held until after the second experiment.

### **4.3 Experiment 2**

This experiment was a speeded lexical decision task. This sort of task taps into a lexical level of processing as opposed to prelexical. In this manner, one can determine if word-final voicing is encoded in the lexical representation of each word.

#### **4.3.1 Materials**

The materials were the unfiltered test items from the first experiment.

#### **4.3.2 Procedure**

Participants were tested individually, with audio presented over stereo headphones (Sony, MDR-V600). They were asked to decide if the word they heard was a real English word or not by pressing the indicated keys on the keyboard, 'yes' for yes, and 'no' for no. Each item was followed by a 1500ms interval of silence during which the participant had to make their decision. After this time, the program continued on to the next word. The order of each item was randomized for each participant. The experiment began with 10 practice trials. In this phase, participants received the feedback 'no response' if they never answered in time, and 'faster' if they did. No feedback was given to indicate the correctness of the response. After the practice trial, time was given for participants to ask questions. They could then go on to the experiment phase, during which, no feedback was given. The experiment took approximately 7 minutes.

#### **4.3.3 Results and discussion**

The break down of error rates for each language group is recorded in (7). A subjects and items anova was conducted to examine if the variance in error rates was significant. For the subjects anova, the within-subjects factors were word type (real or non), voicing (word-finally voiced or voiceless), and place (word-final consonant was labial, coronal, or velar) and the between-subject factor was language. The opposite held for the objects anova.<sup>1</sup> The results are recorded in

---

<sup>1</sup> Two other subjects and items anovas were conducted, one additionally including place and the other manner as a within-subjects factor. In the first case all instances of continuants were removed and in the second all velar and coronal conditions to create more homogenous datasets. As neither attributed to any significant effect ( $F(2, 85) p >$ .

(8). In the subjects anova, the only effects reaching significance were language and word type (real or non). No interactions were significant. In the objects anova, the only effect *not* reaching significance was voicing.

(7) Average Error Rates (%) on LD Task

Group	Real Words			Non-Words			Total
	Vcd	Vcl	Total	Vcd	Vcl	Total	
German	33	14	23	45	71	58	42
English	8	5	7	20	15	18	12

(8) Subjects and Items Anova Results for LD task

Effect	F1	p<0.5?	F2	p<0.5?
Language	1,22.8	yes	1,123	yes
Word Type	1,10.5	yes	1,42.9	yes
Voicing	1,0.84	no	1,2.1	no
Language*Word Type	1,1.68	no	1,7.5	yes
Language*Voicing	1,3.32	no	1,9	yes
Word Type*Voicing	1,2.54	no	1,11.5	yes
Language*Word Type*Voicing	1,1.68	no	1,15	yes

Several observations are apparent from these results. The first, and perhaps most obvious, being that the German English listeners performed significantly worse than the English native listeners. The second is that both groups display above chance (>50%) performance. Both groups performed better with real words than with non-words, with the German English listeners being significantly poor with the non-words. Perhaps more interesting than this interaction is the interaction between language, word type, and voicing. While the English native listeners perform consistently poorer with non-words as well as word-finally voiced words, German English listeners' performance varies. Under the assumption that German English listeners are deaf to word-final voicing, one would expect voicing to have no effect at all. However in the real-word condition German English listeners perform better in the voiceless condition (32% errors for voiced vs. 8% errors for voiceless), while in the non-word condition, they perform better with the voiced (71% vs. 45%). Possible interpretations for this might be that German English listeners are biased towards thinking that words with voiced consonants are English non-words, or that words ending with voiceless consonants are English real words, or both. This in conjunction with the above chance error rate sheds serious doubt on

05, F2(2., 75) p>. 05 and F1(1., 23) p>. 05, F2(1., 40) p>. 05 respectively) these factors were not looked at further and the subjects and items anova reported in this paper contain all of the data).

German English listeners being deaf to English word-final voicing. As a reminder, the question this experiment was answering is: can German English listeners attend to word-final voicing during an online lexical task? The answer is yes, but not well. Even being generous so as to extend the chance range to 45-55%, the participants performed, albeit marginally, above chance.

## 5. General Discussion

The goal of this paper was to experimentally distinguish between structure-adding and structure-changing models of speech perception. Experiments similar to those used by Dupoux and colleagues to examine stress-deafness were performed for this purpose. If a participant was shown to perceive a contrast (experiment 1) but not be able to use it lexically (experiment 2), then structure-adding models of perception are supported. Otherwise, structure-changing models are supported. German listeners of English provided a unique way to test established structure-adding and structure-changing models of second language speech perception. While voicing is contrastive in German, this contrast is neutralized word finally due to obstruent devoicing, i.e., there does not exist a German word in which the last obstruent is pronounced as voiced. Because obstruents are redundantly voiceless in German, the perception model of Dupoux and colleagues suggests that German listeners do not encode word-final voicing, and hence would not be able to use word-final voicing information for lexical decisions even in languages, like English, which display such a contrast. Structure-changing models, however, suggest that because voicing is contrastive in German, German listeners should be able to perceive the contrast and use this perception regardless of position. German English listeners then provided an excellent way to distinguish between the two model types. A disparity between the ability to perceive word-final voicing and being able to use this perception lexically would support structure-adding models, a lack thereof, structure-changing. This was research question number one. There was, however, another reason why German English listeners were chosen. Studies on phonological deafness receive most, if not all, of their support by studies on suprasegmentals. German English listeners provide an opportunity to test the structure-adding deafness theory on segmentals. In the ABX task, German listeners of English were able to perceive English word-final voicing (15% error rate). In the lexical decision task, they were only marginally able to use this perception (42% error rate). This supports structure-adding models. It does not, however, support Dupoux and colleagues model, as German listeners were above chance on lexical decision performance. Shedding further doubt to this theory is German's sensitivity to word-final voicing, as demonstrated by the interaction between word type and voicing.

## References

- Best, Catherine T., Gerald W. McRoberts, and Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of Acoustical Society of America*

109 (2): 775-794.

- Crowther, Court. S., and Virginia Mann. 1992. Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of Acoustical Society of America* 92 (2): 711-722.
- Dehaene-Lambertz, Ghislaine. 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919-924.
- Dupoux, Emmanuel, Pallier, C., Kakehi, K., and Mehler, J. 2001. New evidence for prelexical phonological processing in word recognition. *Language and cognitive processes* 16 (5): 491-505.
- Dupoux, Emmanuel, Christophe Pallier, Núria Sebastián-Gallés, and Jacques Mehler. 1997. A destressing “deafness” in French? *Journal of Memory and Language*, 36, 406-421.
- Dupoux, Emmanuel, and Sharon Peperkamp. 2002. Fossil markers of language development: Phonological ‘deafness’ in adult speech processing. In *Phonetics, phonology, and cognition*, eds. J. Durand and B. Laks, 168-190. New York: Oxford University Press Inc.
- Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete, and Sharon Peperkamp. 2008. Persistent stress deafness: The case of French learners of Spanish. *Cognition* 106 (2): 682-706.
- Gandour, Jack, Donald Wong, and Gary Hutchins. 1998. Pitch processing in the human brain is influenced by language experience. *NeuroReport* 9: 2115-2119.
- Grieser, DiAnne, and Kuhl, Patricia K. 1989. Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology* 25 (4): 577-588.
- Hauser, Marc D., Elissa L. Newport, and Richard N. Aslin. 2001. Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* 78 (3): B53-B64.
- Iverson, Gregory K., and Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12 (3): 369-396.
- Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh'ich Tohkura, Andreas Kettermann, and Claudia Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87: B37B57.
- Jusczyk, Peter W., Paul A. Luce, and Jan Charles-Luce. 1994. Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33: 630-645.
- Klein, Denise, Robert J. Zatorre, Brenda Milner, and Viviane Zhao. 2001. A cross-linguistic pet study of tone perception in Mandarin Chinese and English speakers. *Neuroimage* 13: 646-653.
- Kuhl, Patricia. K. 1991. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50 (2): 93-107.
- Kuhl, Patricia. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences USA* 97 (22): 11850-11857.
- Kuhl, Patricia. K., and James D. Miller. 1975. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* 190 (4209): 69-72.

- Maye, Jessica, Janet F. Werker, and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.
- Newport, Elissa L., Marc D. Hauser, Geertrui Spaepen, and Richard N. Aslin. 2004. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49: 85-117.
- Peperkamp, Sharon, Emmanuel Dupoux, and Sebastián-Gallés, Núria. 1999. Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of Eurospeech '99* 6: 2683-2686.
- Polka, Linda, and Janet F. Werker. 1994. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20 (2): 421-435.
- Raphael, Lawrence J. 1972. Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of Acoustical Society of America* 51: 1296-1303.
- Rivera-Gaxiola, Mariza, Gergely Csibra, Marc H. Johnson, and Annette Karmiloff-Smith. 2000. Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research* 111 (1-2): 13-23.
- Saffran, Jenny R., Richard N. Aslin, and Elissa L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274: 1926-1928.
- Trehub, Sandra E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466-472.
- Werker, Janet F., and Richard C. Tees. 2002. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 25: 121-133.
- Wiese, Richard. 2000. *The phonology of German*. Oxford: Oxford University Press.