

Phonetic Training Significantly Mitigates the Stress 'Deafness' of French Speakers

Angela C. Carpenter

Cognitive & Linguistic Sciences, Wellesley College

106 Central Street, Wellesley, Massachusetts, USA

Tel: 1-781-283-2889 E-mail: acarpent@wellesley.edu

Received: May 23, 2015 Accepted: June 10, 2015 Published: June 30, 2015

doi:10.5296/ijl.v7i3.7661 URL: <http://dx.doi.org/10.5296/ijl.v7i3.7661>

Abstract

Stress 'deafness' refers to the inconsistent perception and/or processing of phonological stress by speakers of fixed stress languages such as French. This paper briefly reports on the results of a study in which French and English participants performed an ABX word stress task, similar to Dupoux et al.'s (1997) Experiment 1. One group of French and a group of English speakers received phonetic training designed to improve perception while two other groups of French and English speakers received no training. The training was an adaptation of the perceptual fading technique, which exposes listeners to stressed syllables that exaggerate the durational correlate of stress, then gradually reduces the durations of subsequent stressed syllables to increase participants' overall ability to accurately perceive stressed syllables. The trained French group performed significantly better than the untrained group with fewer errors and lower response times. As expected there was no difference in accuracy between the trained and untrained English groups. We argue that by exaggerating the duration cue for stress, the phonetic training led to increased overall perception, perhaps even beginning to build an abstract phonological representation of stress that was then carried into the ABX task. Although trained on artificially manipulated stimuli, participants were able to perform well on naturally-produced novel stimuli.

Keywords: Stress 'deafness', French, Phonetic training, Stress, Duration

1. Introduction

Stress ‘deafness’, a term used by Dupoux, Pallier, Sebastian and Mehler (1997), refers to the inconsistent perception and/or processing of phonological stress by speakers of fixed stress languages such as French. By contrast, native speakers of Spanish, in which contrastive stress can occur on different syllables of a word, are significantly more accurate in identifying and discriminating contrastive stress. This phenomenon of stress ‘deafness’ has been observed and studied in various languages including Finnish, Turkish, Hungarian, Polish, and Persian but mostly in French (Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997; Lukyanchenko, Idsardi, & Jiang, 2010; Peperkamp & Dupoux, 2002; Peperkamp, Vendelin, & Dupoux, 2010). Extensive research by Peperkamp, Dupoux and colleagues has shown that stress ‘deafness’ is quite robust and persistent (Dupoux, Sebastien-Galles, Navarrete, & Peperkamp, 2008), and even continues to exist in simultaneous French-Spanish bilinguals (Dupoux, Peperkamp, & Sebastián-Gallés, 2010). Given the persistence of this phenomenon, the question arises, can stress ‘deafness’ be mitigated through phonetic training?

To acquire a phonological distinction in a nonnative language requires that listeners first perceive the distinction, which could then lead to necessary changes in higher-order processing, including building a phonological representation. Perceptual learning, the process through which learners improve their ability to identify and discriminate linguistic sounds through training or experience, aided by selective attention, has been used as a means of improving the acquisition of nonnative segmental contrasts (Pisoni, Lively, & Logan, 1994). Nosofsky's model of selective attention to visual stimuli (Nosofsky, 1986, 1987) has been applied to speech perception and explains the connection between selective attention and speech perception in this way: Requiring the learner to attend to a specific dimension of the stimulus ‘stretches’ the psychological space, such that the perceived distance between the target and other similar sounds is increased (Pisoni et al., 1994). In the study to be discussed here, we hope to demonstrate how selective attention to acoustic cues for stress can stimulate perceptual learning of stress.

Training on specific phonetic cues can help learners of a second language to increase perceptibility of nonnative segmental sounds. Segmental training has included the English /ʊ ~ /Ø/ contrast for French speakers (Jamieson & Morosan, 1986), and most notably, the English /l/ ~ /r/ contrast for Japanese speakers (Best, 1995; Iverson, Hazan, & Bannister, 2005; Lively, Logan, & Pisoni, 1993; Logan, Lively, & Pisoni, 1991; Strange & Dittman, 1984; Takagi, 2002). Training on suprasegmentals has lagged somewhat behind that of well-established phonemic segmental contrasts but work has been done in the areas of tone (So, 2006; Wang, Spence, Jongman, & Sereno, 1999) and stress (Archibald, 1992; Carpenter, 2006, 2010; Yeung & Nazzi, 2014). Wang et al. (1999) report on the training of American English listeners to perceive Mandarin tones more accurately. Adapting the high-variability, multi-talker paradigm developed for segmental training (Lively et al., 1993) to suprasegmental training, Wang et al. (1999) found that learners of Mandarin tone were able to increase their overall tone perception accuracy by a significant 21%. So (2006) compared two different training methods in the acquisition of Mandarin tone by native Cantonese, Japanese and English speakers. She found that Japanese (but not English or Cantonese) learners

improved their perception of Mandarin tone after being trained through either simple feedback, where listeners were just told whether their response was correct or not, or an audio-visual method where specific attention was drawn to a crucial phonetic cue. In a pilot experiment, Carpenter (2006) trained native French speakers to more accurately identify stressed syllables in multisyllabic nonsense words by manipulating the duration, pitch and intensity of stressed syllables to gradually increase listeners' perception. The group that received the training increased their accuracy by 36%. A study with 10-month-old French-learning infants has also shown a positive effect of audiovisual phonetic training to distinguish lexical stress when the words are presented as object labels (Yeung & Nazzi, 2014). In this study one group of French infants were trained on iambic and trochaic disyllabic words that were consistently paired with two objects while a second group heard the same words, but these were inconsistently paired with different objects. Infants were then tested on their ability to generalize the stress pattern to novel words. Infants were able to correctly identify stress, but the group with the inconsistent training was more delayed.

These studies suggest that very young learners appear to have an ability to develop new phonological representations and that adult speakers of English can learn a suprasegmental contrast with phonetic training. However, English speakers already have experience with pitch variation (as in statements versus questions), so perhaps learning a tone distinction is feasible. More research is needed to answer the question of whether phonetic training can increase French speakers' perception of lexical stress, which requires attention to the prosodic features of pitch, duration and intensity. This paper describes such a training and discusses its implications for mitigation of stress 'deafness', and by extension, other suprasegmental features of language, such as tone and prosody.

1.1 Phonetic Correlates of Stress

Stressed syllables are more prominent than the other syllables in a word. The phonetic correlates of this prominence are generally described as being a combination of pitch, duration, intensity and spectral balance (Hayes, 1995; Lieberman, 1960; Sluijter & van Heuven, 1996). Stress in English, for example, is phonetically realized as a combination of duration, intensity and pitch along with vowel reduction of unstressed syllables (Fry, 1955, 1958). For French speakers, duration is a strong cue for stress¹ (DeLattre, 1966; Fant, Kruckenberg, & Nord, 1991). Therefore, in creating stimuli for the training portion of the experiment described in this report we manipulated duration, pitch and intensity along with additional manipulations in duration to increase the perceptibility of stress for French speakers. The paper is organized as follows: Section 2 discusses the experiment including how the stimuli were constructed; Section 3 states the results and the paper concludes in section 4 with a discussion of the results and some implications for second language learning.

¹ Although the primary use of syllable lengthening in French is to demarcate word or phrase endings, French listeners tend to associate duration with prominence.

2. Experiment

The experiment consisted of a training task followed by a discrimination task that involved a stress contrast. The training was an adaptation of the perceptual fading technique where listeners heard stressed syllables with exaggerated durations that were shortened over time. This technique facilitates selective attention to duration. Duration is correlated to prominence in French (DeLattre, 1966; Dell, 1980; Fant et al., 1991), so the duration cue for stress was more heavily manipulated for the training task. Learners thus began the task with tokens that were fairly easy to discern and were then progressively trained to distinguish stress without an exaggerated duration.

The discrimination task replicates and extends the first experiment reported in Dupoux et al. (1997), an ABX design where native speakers of French and Spanish were asked to discriminate nonsense words with a stress contrast. The present experiment tested two groups of native French speakers and two groups of English speakers. One French group received phonetic training while the French control group remained untrained. The prediction was that the trained group would have more correct responses and faster response times on the ABX task than the untrained group. If the trained group performed better on the ABX task that would suggest that the training played a role in increasing their perceptual ability and served to mitigate some of the effects of stress ‘deafness’. Another possibility is that better performance by the trained group could simply reflect a benefit of increased exposure to stress stimuli and not a benefit of the training per se. To rule out that explanation, one group of English speakers was also trained and the other was untrained. English speakers are already sensitive to the correlates of stress since it is unpredictable in English. Thus, training should make no difference to their accuracy and there should be no significant difference between the performances of the two groups of English speakers.

2.1 Methods

2.1.1 Participants

Participants, all paid, were adult native speakers of French in Paris or of English in Massachusetts, USA. There were a total of 35 French (23 F, 12 M) and 35 English participants (23 F, 12 M). Each language group had 17 trained and 18 untrained. French participants were 18 to 37 years of age, $M = 21.61$ ($SD = 5.23$) and English participants were 18 to 38 years of age, $M = 24.21$ ($SD = 6.82$). Fifteen trained participants (7 French and 8 English) were not included as they scored less than 70% correct on the final training block words of 120% duration. None of the participants reported having hearing problems or of being fluent in the other language.

2.1.2 Stimuli

Building the training stimuli. The training stimuli were 3-syllable nonsense words made up of CV and CVC syllables. Onsets were [b, d, g, p, t, k]; codas were [f, ʃ, ʒ, dʒ]; and vowels were [a, i, u]. Syllables were recorded in a sound-attenuated room directly into Praat (Boersma & Weenink, 2013) on a Mac computer by a trained female phonetician who is a native speaker of American English. The sampling rate was 44100 kHz.

Several manipulations were made in Praat to increase the perception of stress. Previous research has shown that English speakers are able to accurately distinguish stressed syllables from unstressed ones when amplitude, pitch and duration are 20% greater in stressed than in unstressed syllables (Carpenter, 2006). Thus, the amplitude of each syllable was equalized to 70 decibels and then increased by a further 20%. Pitch was also increased by 20% and a pitch contour was overlaid on each syllable to approximate the normal variation in pitch based on position of the stressed syllable in a multisyllabic word as observed in pilot experiments conducted with native English speakers (Carpenter, 2006). To train participants on the duration correlate of stress, syllable durations were increased by 80% (longest), 50% (longer) and 20% (long) to create three different sets of lengthened syllables. Syllables with initial stress received a falling contour, those with medial stress had a rising then falling contour and final stress syllables had a rising contour. Following manipulation the syllables were resynthesized using the Pitch Synchronous Overlap Added (PSOLA) algorithm in Praat. To produce the unstressed syllables, the intensity of the original recorded syllables was reduced by 6 dB. These adjustments maximized the perceptual difference between stressed and unstressed syllables. Three-syllable words each had one stressed syllable and two unstressed ones.

Stimuli for the ABX task. The stimuli replicated that of Dupoux, et al. (1997) with twelve triplets of the form 'mepado, me'pado, mepa'do. The words used were: *baveta, bopelo, detoma, lumisa, mepado, metilo, picadu, povami, rimato, someta, tamido, and vasuma*. All stimuli were nonsense words in both French and English. Following Dupoux, et al. (1997) the twelve sets of triplets were recorded by a male and female speaker of Dutch, a language different from both English and French. Speakers were instructed to emphasize the stressed syllable in each word and not to reduce any vowels. The recording was made with Garageband software at a sampling rate of 44100 kHz into a MacBook Pro computer.

Ninety-six ABX trials of triplets were created, with the first two words, A and B, said by the female speaker and the third word, X, said by the male. All three words had the same segmental content. The first two words had stress on different syllables and the third word had the same stress as either the first or the second. In these Dutch recordings stressed syllables averaged 33% higher in pitch, 20% longer in duration and 5.47 dB higher in intensity than unstressed syllables. These differences were significant at $p < .001$ for all three measures. Thus the naturally-produced stimuli recorded by Dutch speakers for the ABX task had phonetic correlates of stress that were similar to the artificially-produced stimuli used in the training task.

Half the trials in the ABX experiment presented an A-B-B sequence while the other half had an A-B-A sequence. The contrasts tested were: 1st vs. 2nd and 2nd vs. 3rd, for example, '*someta~so'meta* and *so'meta~some'ta*. The trials were balanced as to order of presentation of contrast and sequence. The 96 trials were evenly divided into two blocks of 48 trials each and all contrasts were represented evenly in both blocks.

2.1.3 Procedure and Design

Participants in the training condition first heard the words with syllables of longest duration (80% longer than the original syllables), then those of longer duration (50% longer) and finally those with syllables of typically long duration (20% longer than unstressed syllables). They chose the stressed syllable by pressing the key labeled “1”, “2”, or “3” on the keyboard. If they chose correctly the next word was presented, but if they chose incorrectly then the same word was presented again until they chose correctly². Participants heard 60 words of each lengthened syllable duration type, presented in three blocks for a total of 180 training words. After completing the training, which lasted about 15 minutes, they went on to the main experiment.

Ten practice trials with feedback were presented before the main trial blocks. Participants were instructed to listen carefully to groups of three words where the first two words were different and the third was either like the first or the second. They were then to press “1” if the third word sounded like the first, or to press “2” if the third word sounded like the second. When the response was correct they received the visual feedback, “Correcte” or “Correct”. If the response was incorrect the message read, “Votre réponse n’est pas correcte - essayez de nouveau” or “Your answer was not correct – try again.” The trial was then repeated. Participants could not proceed to the next trial until they got each practice trial correct. After the ten practice trials they proceeded to the main ABX task, similar to the practice but without feedback. Each experiment trial presented a triplet, separated by 500 ms. of silence. The next trial began 1000 ms. after a response, or after 4000 ms. if there was no response.

The experiment was built and conducted in SuperLab Pro and all analyses were conducted using SPSS 21.

3. Results

Table 1 lists mean response times and error rates by Language (French or English), Training within the French group, Accent Contrast (1st vs. 2nd and 2nd vs. 3rd) and Response Type (A or B). One accent contrast for the item *tamido* was technically flawed thus all the 1st vs. 2nd accent contrast tokens of *tamido* were omitted from the analyses. Results are detailed in Table 1 showing mean response time in milliseconds, standard error and error rate.

² Some participants chose the same incorrect word multiple times, perhaps indicating that they really had difficulty 'hearing' the stressed syllable.

Table 1. Mean Response Time (RT), Standard Error (SE), and Error Rate (ERR) of ABX Judgments based on Two Accent Contrasts

Language	Response Type	Accent 1 st vs. 2 nd (<i>'bopelo</i> vs. <i>bo'pelo</i>)			Accent 2 nd vs. 3 rd (<i>bo'pelo</i> vs. <i>bope'lo</i>)			Overall Mean		
		Mean RT	SE	ERR	Mean RT	SE	ERR	Mean RT	SE	ERR
French										
Untrained	X = A	1068	43	39%	1040	38	32%	1053	29	36%
	X = B	910	43	26%	901	40	22%	905	29	24%
	Mean	993	31	33%	970	28	27%	981	21	30%
Trained	X = A	858	38	22%	892	36	23%	876	26	23%
	X = B	814	39	20%	775	32	12%	793	25	16%
	Mean	837	27	21%	833	25	18%	835	18	19%
All French	X = A	966	29	31%	968	26	28%	967	20	29%
	X = B	864	29	23%	840	26	17%	851	19	20%
	Mean	917	21	27%	904	19	22%	910	14	25%
English										
Untrained	X = A	880	38	25%	934	38	15%	908	27	20%
	X = B	828	42	16%	656	27	8%	737	24	12%
	Mean	856	28	20%	795	24	12%	824	18	16%
Trained	X = A	1016	48	21%	1012	46	13%	1014	33	17%
	X = B	961	54	14%	749	40	6%	847	33	10%
	Mean	990	36	18%	880	31	9%	932	24	13%
All English	X = A	946	30	23%	973	30	14%	959	21	18%
	X = B	892	34	15%	701	24	7%	790	20	11%
	Mean	919	22	19%	834	20	11%	877	15	15%

Error rate (ERR) in Table 1 refers to the percentage of incorrect responses. Response Type refers to the position of the final stimulus (the X) within the ABX design. For example, if participants heard *'someta~so'meta~'someta*, the correct response for the X stimulus (*'someta*) would be A which is the same as the first stimulus. If they were presented with *so'meta~some'ta~some'ta* the correct response would be B because the X stimulus *some'ta* copies the second stimulus.

Overall, English participants were more accurate than French (85% vs. 75%) and trained participants performed better than untrained (84% vs. 77%). Trained French participants performed better than untrained French by making fewer errors (19% vs. 30%) and having a quicker response time (835 ms. vs. 981 ms.). Trained English performed slightly better than untrained English in error rates (13% vs. 16%) but were slower in reaction times than the untrained (trained – 824 ms., untrained – 932 ms.). Figure 1 shows the proportion correct for trained and untrained French and English.

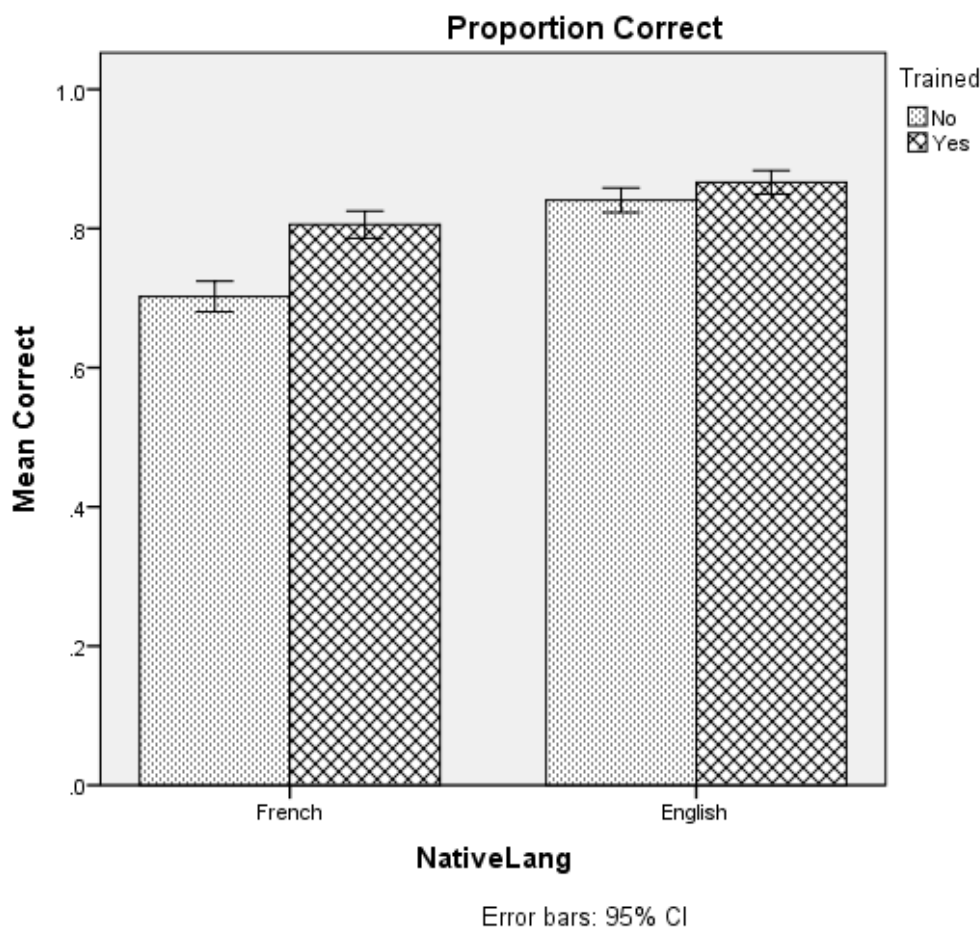


Figure 1. Proportion correct of untrained and untrained French and English participants

A binomial logistic mixed model regression was used to analyze the probability of participants choosing the correctly stressed word given the choice of a correct and an incorrect response in each trial. The dependent variable was the number of correct answers and the independent variables were native language (French or English) and training (Yes or No). Table 2 presents the results of the mixed model including all participants.

Table 2. Mixed Model Regression Analysis of Correct Responses by Native Language and Training

	Estimate	Std. Error	z	p
(Intercept)	1.6258	0.1158	14.042	<.0001
Native Language	0.35831	0.0967	3.704	.0002**
Trained	-0.22630	0.0973	-2.326	.020*
Native Language X Trained	0.08414	0.09631	0.874	.382

* significant at the $p < .05$ level, ** significant at the $p < .001$ level

English participants outperformed French participants ($t = 3.70$, $p < .001$) and the untrained

trials were less accurate than the trained trials ($t = -2.32, p < .05$). There was no interaction ($t < 1, p > .38$).

We were particularly interested in how the separate language groups performed based on whether or not they had received training to improve perception of stressed syllables. The sections below report on the planned comparisons of French and English performance.

French: Trained vs. Untrained. Trained French participants performed significantly better than untrained French by making fewer errors (19% vs. 30%) as seen in the regression analysis in Table 3.

Table 3. French speakers: Trained vs. Untrained

	Estimate	Std. Error	z	p
(Intercept)	1.2615	0.1387	9.094	<.0001
Trained	-0.3099	0.1218	-2.545	0.0190*

* significant at the $p < .05$ level

English: Trained vs. Untrained. There was no significant difference in proportions correct between trained and untrained English participants as seen in Table 4.

Table 4. English speakers: Trained vs. Untrained

	Estimate	Std. Error	z	p
(Intercept)	1.9952	0.1631	12.229	<.0001
Trained	-0.1414	0.1508	-0.938	0.348

English speakers, trained and untrained, were able to accurately identify the stressed pairs in the ABX task. This result is as expected because English speakers are already well-acquainted with stress differences from their native language.

French vs. English. Among participants who did not receive any training, English speakers were significantly more accurate than French, ($t = 3.51, p < .001$). This result is similar to that of Dupoux et al.'s (1997) finding where Spanish speakers were significantly more accurate than French.

Table 5. French and English untrained

	Estimate	Std. Error	z	p
(Intercept)	1.3949	0.1351	10.32	<.0001
Native Language	0.4421	0.1260	3.51	.0004**

** significant at $p < .001$ level

Among those that were trained, English speakers were marginally more accurate than French speakers ($t = 1.89, z = 1.89, p = .058$).

An ANOVA of reaction times was performed with time as the dependent variable and Native Language and Training as independent variables. There was no significant difference between

the two groups based on language ($p = .350$) and training ($p = .754$), however there was an interaction between language and training ($p < .001$). This interaction was a result of the differing direction of the two groups: Among the English group, trained participants were slower than untrained ones (932 ms. vs. 824 ms., $F(1, 3218) = 21.708$, $p < .001$, $\eta^2 = .007$), but among the French, trained French participants were faster than those untrained (835 ms. vs. 981 ms., $F(1, 3218) = 31.72$, $p < .001$, $\eta^2 = .01$).

4. Discussion and Conclusion

This study sought to replicate and extend a previous study on stress 'deafness' done by Dupoux et al. (1997) in which French and Spanish speakers were asked to perform an ABX task. Participants were instructed to choose the pair of nonsense words within a triplet that were all matched segmentally but within which one pair also matched prosodically. Dupoux et al. found that French speakers had much more difficulty in performing the task than did Spanish speakers, as evidenced by the significantly different error rates of 19% for French and 4% for Spanish. Mean response times showed that French participants were slower than Spanish participants (1236 ms. vs. 1142 ms.), although this result was not significant. Following Dupoux et al.'s design, this study compared French and English participants' performance in the same task. However, there were two groups in each language, one which received specialized training to help them better perceive stressed syllables and another that did not receive the training. French speakers with no training made significantly more errors and had slower response times than did native English speakers with no special training. These results confirm those of the many studies which demonstrate that speakers of a fixed stress language have difficulty in perceiving and processing word-level contrastive stress (Dupoux et al., 1997; Dupoux, Peperkamp, & Sebastián-Gallés, 2001; Dupoux et al., 2008; Peperkamp & Dupoux, 2002; Peperkamp, Dupoux, & Sebastián-Gallés, 1999). While the trained English group was slightly more accurate than the untrained group, the difference was not significant, thus reinforcing the idea that training, and not mere increased exposure, contributed to the better success rate of the trained French group. Interestingly, English speakers in this experiment were less accurate in the task than were Spanish speakers in Dupoux et al. (1997). This result is not surprising however since English speakers use different cues than Spanish speakers to signal stressed and unstressed syllables in a multisyllabic word. In addition to the acoustic cues of duration, pitch and intensity to identify stressed syllables, English uses reduced vowels in unstressed syllables (Chomsky & Halle, 1968; Ladefoged, 1975), where Spanish does not have reduced vowels (Best & Tyler, 2007; DeLattre, 1966). The unstressed vowels in this experiment were not reduced, therefore omitting a strong cue for English-speaking participants.

A growing body of research into the typology of stress 'deafness' has led to models that capture some of the linguistic factors that produce this type of phonological 'deafness'. The Stress Deafness Model (SDM) (Dupoux et al., 1997; Dupoux & Peperkamp, 2002; Peperkamp & Dupoux, 2002) predicts that languages in which stress is surface-evident will exhibit stress 'deafness' to some degree, mitigated by the presence of lexical exceptions in the language. Languages with lexical exceptions produce less 'deafness' than exceptionless ones. The Stress Typology Model (STM) (Altmann, 2006; Altmann & Vogel, 2002) presents a

typology based on a binary branching system of language classification which includes stress and non-stress languages. Like the SDM, this model also predicts that speakers of a language with predictable stress would perform more poorly in identifying stressed syllables than those who speak languages with unpredictable stress. The STM differs from the SDM, though, in that it predicts that speakers of non-stress languages, those with tone and pitch accents, should be able to accurately identify stressed syllables. The result of the experiment in the current study is consistent with the STM and SDM models' prediction that speakers of languages with unpredictable stress are already well attuned to stressed syllables. Further training makes little difference in their ability. However, while both models address the perception of stress, neither of them predicts that perception could be changed, by training or otherwise. Perhaps future extension of these models should take into account a possible effect of phonetic training to mitigate the effects of stress 'deafness'.

While selective attention was incorporated in the experimental design to facilitate perceptual learning, did any learning that may have occurred lead to building a new or modified phonological representation of stress? There are two facts that suggest that learners may have begun the process of building an abstract representation of stress. First, the composite nature of the training words was different from the test words. The training words were artificially manipulated, but learners were presented with *naturally-produced* nonsense words in the ABX task. The trained group could not carry over their short-term auditory memory of the exact acoustic cues in the artificially-produced training words to the naturally-produced test words as short term store is of short duration, generally less than 30 seconds (Atkinson & Shiffrin, 1968). Further, stressed syllables in the training words were made up of CV and CVC syllables while the stressed syllables in the test were only CV, thus increasing the difference between the training and test tasks. Second, trained learners had to identify stress across two different accents. They were able to carry over their increased perception of the stress cues from English-accented words to Dutch-accented ones across different voices. Since the trained French participants were significantly more accurate than the untrained, this suggests that a nascent phonological representation began to be built to organize the acoustic cues. The research done here, while preliminary, suggests that focusing attention on a relevant cue that is discernible to the learner could help to bridge the gap between listeners' discernment of a cue and their development of a more durable phonological representation.

How exactly suprasegmental effects such as stress are encoded in the brain is not presently known. However, it is likely that differences will exist between language groups and the target languages in terms of what is needed for increased perception that leads to abstract encoding. For French speakers, the training stimuli increased all stress cues including intensity, pitch and duration and then duration, which is already salient to French speakers, was further exaggerated. While duration was the important cue, to state definitely that the duration cue was solely responsible for better performance would require further research, including manipulating pitch and intensity and comparing the differences. Using phonetically-varied stimuli might also clarify the issue of building a phonological representation.

The results of this study demonstrate that it is possible to mitigate the effects of stress

'deafness' through specialized training and selective attention to a perceptible cue. Further research needs to be done to explore if phonetic effects last over a longer period of time, and whether such training also improves production. Nevertheless, this brief study demonstrates the possibilities of phonetic training to aid in the acquisition of suprasegmental features. This result is an encouraging one for L2 researchers, teachers and learners as it shows promise for the potential of training learners to better discriminate and acquire additional prosodic contrasts, including phonemic tone and other forms of word prosody.

Acknowledgements

This work was supported by a Woodrow Wilson Career Enhancement Fellowship for Junior Faculty and Wellesley College. I would like to thank Andries Coetzee, Pierre Hallé, Andrea Levitt, Joe Pater, and the anonymous reviewers for their helpful comments and suggestions. Thanks also to Ellen Shlasko and Justine Menard for help with recruiting participants in Paris.

References

- Altmann, H. (2006). *The perception and production of second language stress: A cross-linguistic experimental study*. (Ph.D.), University of Delaware.
- Altmann, H., & Vogel, I. (2002). *L2 acquisition of stress: the role of L1*. Paper presented at the DGfS Annual Meeting "Multilingualism Today", Mannheim, Germany.
- Archibald, J. (1992). Transfer of L1 parameter settings: Some empirical evidence from Polish metrics. *Canadian Journal of Linguistics*, 37, 301-339.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation* (Vol. 2, pp. 89-195). New York: Academic Press. [http://dx.doi.org/10.1016/S0079-7421\(08\)60422-3](http://dx.doi.org/10.1016/S0079-7421(08)60422-3)
- Best, C. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167-200). Baltimore: York.
- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & Munro M. (Ed.), *Language Experience in Second language Speech Learning. In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/llt.17.07bes>
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.2.15, retrieved 10 January 2013 from <http://www.praat.org/>
- Carpenter, A. (2006). *Acquisition of a natural versus an unnatural stress system*. (Ph.D.), University of Massachusetts Amherst, Amherst, MA.
- Carpenter, A. (2010). A naturalness bias in learning stress. *Phonology*, 27(3), 345-392.
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

DeLattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics in Language Teaching*, 4(3), 183-198. <http://dx.doi.org/10.1515/iral.1966.4.1-4.183>

Dell, F. (1980). *Generative phonology and French phonology*. Cambridge, UK: Cambridge University Press.

Dupoux, E., & Peperkamp, S. (2002). Fossil markers of language development: Phonological 'deafnesses' in adult speech processing. In J. Durand & B. Laks (Eds.), *Phonetics, Phonology and Cognition* (pp. 168-190). Oxford: Oxford University Press.

Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language*, 36, 406-421. <http://dx.doi.org/10.1006/jmla.1996.2500>

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America*, 110, 1606-1618. <http://dx.doi.org/10.1121/1.1380437>

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress 'deafness' in simultaneous French-Spanish bilinguals. *Cognition*, 114, 266-275. <http://dx.doi.org/10.1016/j.cognition.2009.10.001>

Dupoux, E., Sebastien-Galles, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, 106(2), 682-706. <http://dx.doi.org/10.1016/j.cognition.2007.04.001>

Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French and English [July-Oct.]. *Journal of Phonetics*, 19, 351-365.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *JASA*, 27, 765-768. <http://dx.doi.org/10.1121/1.1908022>

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 126-152.

Hayes, B. (1995). *Metrical Stress Theory*. Chicago: The University of Chicago Press.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America*, 118(5), 3267-3278. <http://dx.doi.org/10.1121/1.2062307>

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English contrast by francophones. *Perception and Psychophysics*, 40, 205-215. <http://dx.doi.org/10.3758/BF03211500>

Ladefoged, P. (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, 32(4), 451-454.

<http://dx.doi.org/10.1121/1.1908095>

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/ II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94(3), 1242-1255.

<http://dx.doi.org/10.1121/1.408177>

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89(2), 874-886.

<http://dx.doi.org/10.1121/1.1894649>

Lukyanchenko, A., Idsardi, W., & Jiang, N. (2010). *Opening your Ears: The role of L1 in processing of nonnative prosodic contrasts*. Paper presented at the 2010 Second Language Research Forum.

Nosofsky, R. M. (1986). Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology: General*, 115(1), 39-57.

<http://dx.doi.org/10.1037/0096-3445.115.1.39>

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(1), 87-108. <http://dx.doi.org/10.1037/0278-7393.13.1.87>

Peperkamp, S., & Dupoux, E. (2002). *A typological study of stress 'deafness'* (2002 ed. Vol. 7). Berlin: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110197105.203>

Peperkamp, S., Dupoux, E., & Sebastián-Gallés, N. (1999). *Perception of stress by French, Spanish and bilingual subjects*. Paper presented at the EuroSpeech '99.

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, 38, 422-430.

<http://dx.doi.org/10.1016/j.wocn.2010.04.001>

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: implications for theories of speech perception. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words* (pp. 121-166). Cambridge, MA: MIT Press.

Sluifster, A. M. C., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471-2485.

<http://dx.doi.org/10.1121/1.417955>

So, C. K. L. (2006). *Effects of L1 prosodic background and AV training on learning mandarin tones by speakers of Cantonese, Japanese, and English.*, Simon Fraser University.

Strange, W., & Dittman, S. (1984). Effects of discrimination training on the perception of /r-/l/ by Japanese adults learning English. *Perception and Psychophysics*, 36, 131-145. <http://dx.doi.org/10.3758/BF03202673>

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight

case studies. *Journal of the Acoustical Society of America*, 111(6), 2887-2896. <http://dx.doi.org/10.1121/1.1480418>

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, 106(6), 3649-3658. <http://dx.doi.org/10.1121/1.428217>

Yeung, H. H., & Nazzi, T. (2014). Object labeling influences infant phonetic learning and generalization. *Cognition*, 132(2), 151-163. <http://dx.doi.org/10.1016/j.cognition.2014.04.001>

Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).