Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Phonetic training makes word learning easier

Permalink
https://escholarship.org/uc/item/8q76w1sc

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN
1069-7977

Authors

Prefors, Amy
Dunbar, David

Publication Date

2010

Phonetic training makes word learning easier

Amy Perfors (amy.perfors@adelaide.edu.au)
Department of Psychology, University of Adelaide

David Dunbar (david.dunbar@student.adelaide.edu.au)
Department of Psychology, University of Adelaide

Abstract

Motivated by the idea that differences between adult and child language learners may stem in part from initially minor differences (such as in phonetic perception) that cascade throughout other aspects of language learning, we explored to what extent training adults on a novel phonetic contrast results in improved learning of words that incorporate that contrast. Results indicate that distributional training on a novel phonetic contrast improves word learning as well as the ability to discriminate a related contrast. We discuss implications for how adults' phonological abilities in affect other aspects of language learning, and also for understanding the effectiveness of different phonetic training regimes.

Keywords: language acquisition; phonetic learning; second language learning

Introduction

Children and adults differ both qualitatively and quantitatively in their ability to acquire a new language. Adults have difficulty with many aspects of language acquisition, from phonetic perception (Werker & Tees, 1984; Werker & Lalonde, 1988; Kuhl, 2004) to language processing (Clahsen & Felser, 2006) to certain aspects of syntax (e.g., Johnson & Newport, 1989; Birdsong, 2006). Scientists have proposed many theories to account for the difference between children and adults; these theories differ in both the degree and type of contribution made by pre-existing language-specific biases. Although nearly everyone agrees that (due to the inherent logical problem of induction posed by language learning) some bias must be necessary to explain successful language acquisition, explanations about the nature of the bias – and the difference between children and adults – vary considerably.

Some argue that there is a fundamental difference between first and second language acquisition: that acquisition in children is guided by an innate Universal Grammar and language-specific acquisition procedures, but that adult acquisition is directed by more domain-general learning mechanisms (e.g., Bley-Vroman, 1990). There are many other possibilities, however, since children and adults differ profoundly in their cognitive capabilities and typical linguistic input. Children have significantly poorer cognitive skills, including memory and processing speed; perhaps these differences aid children to learn language by enabling them to isolate and analyze components of a linguistic stimulus (Newport, 1988) or to over-regularize inconsistent input (Hudson Kam & Newport, 2005; Singleton & Newport, 2004). Another possibility is that learning a second language is made more difficult due to interference from the first language; indeed, the evidence that experience with a first language influences acquisition of a second is extensive (e.g., Mayberry, 1993; Iverson et al., 2003; Tan, 2003; Weber & Cutler, 2003; Hernandez, Li, & MacWhinney, 2005). This explanation overlaps considerably with the related point that adult brains are in many ways less plastic, and therefore less malleable in response to novel input (Elman et al., 1996; MacWhinney, 2005). Other explanations suggest that adults and children differ in their style of learning (Ullman, 2004) as well as the nature of the social support (Snow, 1999) and linguistic input (Fernald & Simon, 1984) they receive. Of course, many of these possibilities may be true simultaneously.

This work investigates yet another possibility – that small differences in children's abilities along one dimension or aspect of language can have cascading effects, resulting in larger differences in other aspects of language. These initial minor differences might be due to language-specific skills that naturally decay over time, or could be due to domain-general changes in the underlying cognitive abilities that subserve them. Key to this idea is the notion that, because language is such an intertwined, multi-dependent system, small differences in one aspect of language can be steadily amplified when it comes to the acquisition of other aspects. This idea is similar to the neo-constructivist view of Karmiloff-Smith (1998): both suggest that differences in eventual linguistic performance may derive from cascading effects that result from variation in more basic skills. That view focuses on abnormal development in children, however. Our work is motivated by an extension of this viewpoint: the notion that some of the well-attested differences between child and adult learners may result from the more minor, lower-level differences between adults and children. To investigate this, we begin by identifying aspects of language acquisition where one might expect to see cascading effects, and investigate whether performance in one improves performance in the other.

What minor difference between adults and children might have significant cascading effects onto other aspects of language? One possibility derives from children's well-attested superior phonological processing and perception abilities. Young infants can distinguish between phonemes in all natural languages, but lose that ability by the age of 10-12 months if they have not received sufficient linguistic input for a language containing that phoneme (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Werker & Tees, 1984; Kuhl, 2004). Adults who begin acquisition of a language later in life, even after decades of experience using the language, show phonological deficits in perception, production, and processing (e.g., Flege, 1995; Pallier, Colomé, & Sebastián-Gallés, 2001; Sebastián-Gallés & Soto-Faraco, 1999).

Moreover, it is quite difficult to train adults to learn a phonetic contrast that does not exist in their native language. Various training regimes exist; some rely on implicit learning of the phonemic categories based on distributional information (Maye & Gerken, 2001, 2002; Shea & Curtin, 2005; Hayes-Harb, 2007), while in others explicit feedback is given (Jamieson & Morosan, 1989; Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002). Although it is possible to train adults to discern non-native phonetic contrasts, the resulting phonetic representations are often fragile. For instance, when trained through implicit distributional learning, adults show little ability to generalize their knowledge to other non-native contrasts that differ along an analogous phonetic feature (Maye & Gerken, 2001), even though infants are able to do so (Maye, Weiss, & Aslin, 2008).

Why might difficulties in phoneme perception be responsible for adults' relatively poor performance on other aspects of language? It is well-known that adults have difficulty rapidly processing fluent speech in their second language (e.g., Guillelmon & Grosjean, 2001; Clahsen & Felser, 2006), which may be in part due to difficulty in perceiving and representing the phonemes that make up that speech. Difficulties in rapid processing could lead to difficulties in segmenting words and mapping those words onto their correct referents; difficulties in identifying words – particularly function words, which are generally shorter and more phonologically impoverished than content words – might result in more difficulty identifying the appropriate parse for sentences and therefore the correct underlying grammatical structure. Consistent with this, phonological working memory is correlated with second language skills in adults (e.g., Perani, 2005), and speech processing efficiency is related to other aspects of linguistic competence in children (Tsao, Liu, & Kuhl, 2004; Fernald, Perfors, & Marchman, 2006). Empirical evidence reveals that knowledge of lower-level aspects of language (such as phonological perception or statistical segmentation) can help in the acquisition of more complex linguistic phenomena (Werker & Yeung, 2005; Mirman, Magnuson, Graf Estes, & Dixon, 2008). Recent computational work suggests that word learning and phonetic category learning are more effective when occurring simultaneously (Feldman & Griffiths, 2009), and that knowledge of phoneme distributions may aid in speech segmentation and identification of lexical categories (Christiansen, Onnis, & Hockema, 2009). However, there is no work we are aware of that explores whether the ability to recognize a phonetic contrast assists adults in other areas of language.

The work here addresses that issue. We train adult learners to perceive a non-native phonetic contrast and then evaluate how this affects their ability to learn novel words containing the phonetic contrast in question. Our results are relevant not only to the possibility that deficits in phonetic skills may have cascading effects through other aspects of language; they are also relevant to the question of how generalizable adult phonetic learning is. As mentioned previously, existing work suggests that although adults can be trained to distinguish novel contrasts, this ability is fragile, and they have difficulty generalizing that contrast to analogous contrasts (Maye & Gerken, 2001). However, this work used synthesized stimuli not found in any natural language, and training included many filler items, so that there was effectively less than five minutes of exposure to the phonemes of interest. Would adults be able to generalize with more exposure or on a more naturally-produced contrast? In other training regimes adults show robust differences in both perception and production of a novel contrast (Lively, Logan, & Pisoni, 1993; Bradlow et al., 1999; McCandliss et al., 2002), but these regimes differed in many ways from Maye and Gerken (2001): they were significantly longer, used more natural stimuli, and involved explicit training with feedback, among other differences. Most importantly, most of these studies did not evaluate generalization to a novel but similar phonetic contrast. Among those that did, generalization to the novel contrast was successful, but the training paradigms involved giving explicit feedback rather than distributional training (e.g., McClaskey, Pisoni, & Carrell, 1983; Wang, Spence, Jongman, & Sereno, 1999). It is therefore unclear whether the limited generalizability observed in Maye and Gerken (2001) is due some inherent inability to generalize based on distributional information, or is due to other details in the training regime. In this work, we incorporate an implicit distributional training regime similar to that of Maye and Gerken (2001), but one of longer duration and with more natural stimuli. Do these changes in training result in improved generalizability, both in terms of novel but similar phonetic contrasts, but also in terms of the ability to use the new phonetic categories when learning new words?

Method

We trained 61 participants recruited from the student population[1] at the University of Adelaide on two tasks: a phonetic training task and a word-learning task. Participants[2] were randomly assigned to either a CONTROL or a TRAINED condition, which differed in terms of the nature of the phonetic training given.

Task 1: Phonetic learning

Training. The first task consisted of phonetic training based on distributional learning, similar to the task in Maye and Gerken (2001). Subjects in the TRAINED condition were trained on the unaspirated velar plosive voiced/voiceless contrast (/g/-/k/), which occurs in languages such as Hindi but not in English (both phonemes sound like a "g" to an English

---

[1] No participants were native speakers of a language with the phonetic contrast we sought to train. 52 were native English speakers. To ensure that native language was not a factor, we performed all analyses on the full population as well as the English speakers only. Results were identical, so we report the full population results.

[2] Of the original 61 subjects, 9 were excluded from the final analysis (5 due to technical difficulties, 1 for failure to follow instructions and 3 who performed at chance levels on the control task, indicating inattention). This left 25 participants in the CONTROL condition and 27 in the TRAINED condition.

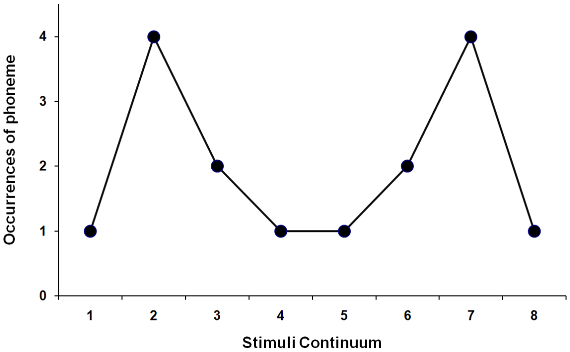

Figure 1: Distribution of stimuli used in phonetic training, defined along a continuum based on VOT. Tokens 2 and 7 occurred four times as often as tokens 1 and 8.

speaker). The /g/ and /k/ phonemes differ in terms of voice-onset time (VOT), such that /g/ contains a pre-voicing component while /k/ does not. It is therefore possible to gradually convert /g/ tokens into /k/ by successively removing parts of the pre-voicing component. Doing so yields a continuum of eight tokens from /g/ to /k/, separated by an average of 17ms in VOT from each other, and identical to each other except for the pre-voicing. As in Maye and Gerken (2001), we presented subjects with a bimodal distribution of these phonemes, as illustrated in Figure 1; thus, some tokens (e.g., 2 and 7) occurred four times as often as others (e.g., 1 and 8). Stimuli were recorded from a male native speaker of Hindi and edited using Praat phonetics software. Each of the phonemes occurred in one of three vowel contexts (/a/, /i/, and /u/).

In order to control for time spent listening to speech sounds across groups, subjects in the CONTROL condition also listened to a distribution of phonemes. However, they heard tokens from a phonemic contrast they could already recognize: the dental plosive aspirated/unaspirated voiced/voiceless contrast (/d/-/t^h/, which sound like "d" and "t" respectively to a native English speaker). As before, these phonemes were used to create a continuum of eight tokens extending from /d/ to /t^h/. Since these phonemes differ along aspiration as well as voicing, the tokens were created by removing voicing and then adding aspiration in continuous steps.

In both conditions, participants listened to a total of 912 tokens presented in random order and separated by 250 ms each, for a total of approximately 11 minutes of exposure to the sounds. During stimulus presentation the participants were told not to speak or read, but also that they need not consciously concentrate on the sounds. To alleviate boredom, they were allowed to doodle while listening.

Testing. Discrimination of the phonetic contrast was tested by presenting participants in both conditions with trials in which they heard three phonemes, two of which were identical. They were asked to press a button indicating whether the third phoneme they heard was the same as the first or the second (the distribution of correct answers was balanced across trials). There were three kinds of trials, defined by the nature of the phonemes tested. On *control* trials, the phonemes already existed in English (/d/ and /t^h/). On the *trained* trials, the phonemes were the ones that the TRAINED group had been trained on (/g/ and /k/). Finally, on the *untrained* trials, subjects were presented with a phonetic contrast that also does not exist in English and that is also defined by voice onset time, but differs in place of articulation – the unaspirated bilabial plosive voiced/voiceless contrast (/b/ and /p/, both of which sound like "b" to an English speaker).[3] The *untrained* trials enabled us to evaluate whether our subjects could generalize any learning to similar phonemes that differed on the same feature. There were 12 test trials for each contrast, totaling 36 testing trials in all; no feedback was given, and the order of all test trials was randomized.

Task 2: Word learning

Training. The word learning task was a standard task in which participants were presented with 12 different image types distributed over three stages of 36 trials each, making 108 trials in all. One each trial, an image was paired with a word, and the participants were instructed to try to learn the word-picture mapping. Words consisted of minimal pairs differing in initial position on each of the contrasts: *trained*: [g]ipur, [k]ipur, [g]anug, and [k]anug; *control*: [d]ipur, [t^h]ipur, [d]anug, and [t^h]anug; and *untrained*: [b]ipur, [p]ipur, [b]anug, and [p]anug. To ensure that the words differed only in the initial sound, words were created by splicing the same stem (-anug or -ipur) to the initial phonemes. The images corresponded to some of the earliest words spoken by children,[4] and were thus presumed to be highly familiar to all participants. The specific image-word pairing was randomized for each participant. The order of presentation of images was also random, with the constraint that each word-image pair was presented three times during each stage.

Testing. There were three testing sessions of 12 trials each, occurring after each stage. During each test trial, one of the 12 images was presented and participants heard two minimal pairs differing along the contrast in question (*trained*, *untrained*, or *control*). Thus, a participant might see a picture of a cat and hear [b]ipur followed by [p]ipur. Their task was to indicate whether the first or second word they heard was correct. No feedback was given.

Results

Task 1: Phonetic learning

Phonetic learning was evaluated by comparing performance on the phonetic test. As Figure 2 illustrates, participants in the TRAINED condition outperformed those in the CONTROL

---

[3] For *trained* and *control* trials, the exemplar tested on corresponded to token 1 and 8 from each continuum. Due to a coding error, the exemplar in the UNTRAINED trials corresponded to tokens 2 and 7 rather than 1 and 8. If anything, this is a more stringent test of generalization, but also means that it is more difficult to compare performance on the UNTRAINED trials to the other two. We discuss the implications of this in subsequent sections.

[4] They consisted of images of babies, balls, books, cats, chairs, birds, beds, cars, cookies, cups, dogs, and shoes.

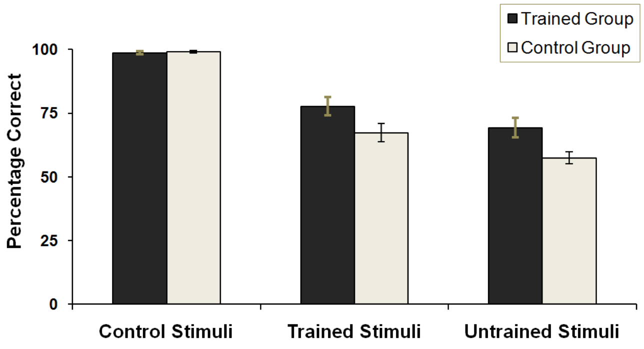

Figure 2: Phoneme discrimination test results. Participants who received distributional training outperformed participants in the CONTROL condition, but all participants performed above chance on all stimuli, suggesting that the test itself may have trained them. Error bars reflect standard error.

condition on both the *trained* and *untrained* stimuli.[5]

Interestingly, participants in both conditions performed above chance on the *trained* and *untrained* stimuli.[6] This suggests that the phonetic testing itself may have trained the participants in the CONTROL condition, which is not an unreasonable suggestion since it closely corresponds to the "prototype" training employed by Jamieson and Morosan (1989) or the "two-seven" condition of Hayes-Harb (2007). To evaluate to what extent such training occurred, we split scores on the phonetic test in half and compared performance on the first six test trials for each stimulus type with performance on the final six test trials of each. As Figure 3 indicates, both groups improved significantly over the course of the test.[7] There was no difference between the CONTROL group's performance in the final half of testing and the TRAINED group's performance in the first half: in other words, training during testing was so effective that it resulted in performance equivalent to having listened to distributional information for over 10 minutes.

It is also evident that performance on the *trained* stimuli was superior to performance on the *untrained* stimuli. This is true even for participants in the CONTROL condition, for whom there should have been no difference between the two types of stimuli (since they had heard neither before). This is probably an artifact of the coding error described earlier in which the *untrained* test stimuli consisted of tokens 2 and 7, rather than tokens 1 and 8 as for the *trained* stimuli. The *trained* stimuli were therefore probably both more effective at teaching participants the contrast, and also easier to differentiate (and hence get correct on the test). Consistent with the hypothesis that this was a training effect, analysis of the first trial of testing reveals that participants in the CONTROL condition performed equally, no better than chance, on both *trained* and *untrained* stimuli. In any case, the important finding – that subjects were able to generalize their phonetic learning to an untrained but related contrast – is unaffected by this detail.

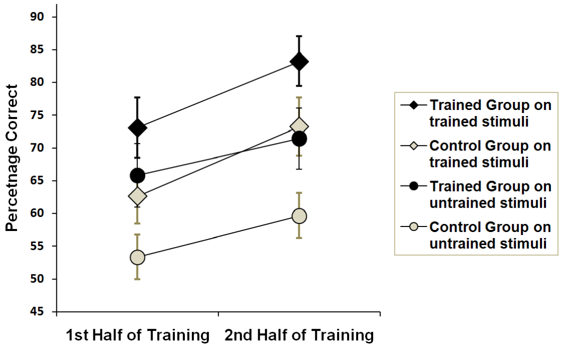

Figure 3: Were participants trained over the course of phonetic testing? Performance on the first half of testing is compared to performance on the second half. Both the TRAINED and CONTROL groups performed significantly better over the course of testing on the *trained* stimuli. While there was a positive trend, the difference in performance on the *untrained* stimuli for either group across the two halves of testing was not significant. The differential effects on *trained* and *untrained* stimuli is probably because the *trained* stimuli were easier to discriminate (tokens 1 and 8) than the *untrained* stimuli (tokens 2 and 7).

Task 2: Word learning

Are participants able to generalize their phonetic discrimination abilities to a new task (word learning), as well as a new contrast? If the phonetic representations acquired are fragile enough, it is possible that they might not, since word learning incorporates many skills: hearing and identifying the phoneme in the context of an entire word; mapping that word onto an image; and doing so while simultaneously trying to learn other word-image mappings. If the task is difficult enough and the representation weak enough, one might expect that it would not transfer.

To answer this question we compared overall performance on the word-learning task, the results of which are shown in Figure 4. As one would expect, participants in both groups were able to identify the *control* words above chance. The TRAINED group performed above chance on the *trained* words, which began with the sound they were trained on; however, they performed at chance on the *untrained* words.[8] By contrast, the CONTROL group was unable to distinguish words beginning with any of the unfamiliar phonemes above chance. There was no difference in performance over the

---

[5] For *trained:* $t(50) = 2.11, p = 0.04$, *untrained:* $t(39) = 2.68, p = 0.011$, both two-tailed. Note that the degrees of freedom for the *untrained* trials were adjusted from 50 to 39; this was because Levene's test for equality of variance indicated unequal variance.

[6] TRAINED group on *trained* stimuli: $t(24), p < 0.001$; on *untrained* stimuli: $t(24) = 5.03, p < 0.001$; CONTROL group on *trained* stimuli: $t(26), p < 0.001$; on *untrained* stimuli: $t(26), p < 0.01$.

[7] Difference between the first and second half of the test trials for the CONTROL participants: $t(26) = 1.87, p = 0.036$; for the TRAINED participants: $t(24) = 2.12, p = 0.022$, both one-tailed.

[8] Differences from chance (50%) performance for the TRAINED group: on words with the *control* contrast: $t(24) = 8.118, p < 0.001$; on words with the *trained* contrast: $t(24) = 2.941, p = 0.007$; on words with the *untrained* contrast: $t(24) = 0.282, p = 0.781$. For the CONTROL group: on words with the *control* contrast: $t(26) = 7.710, p < 0.001$; on words with the *trained* contrast: $t(26) = -0.090, p = 0.929$; on words with the *untrained* contrast: $t(26) = 0.991, p = 0.331$. All tests are two-tailed.

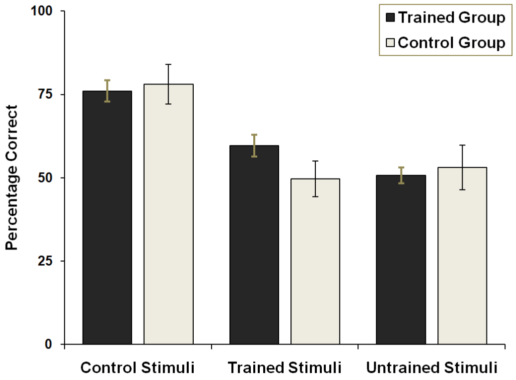

Figure 4: Word learning results. Participants in both groups were able to identify the correct words for the *control* stimuli above chance. The TRAINED group was above chance on words beginning with the sound they were trained on, but not on the related untrained sound. The CONTROL group, which was not trained on any phonemes, was unable to learn words beginning with both the *trained* and *untrained* sounds.

course of the word-learning task for any condition on any stimuli, suggesting that the task did not itself train phoneme discrimination.

Discussion

Motivated by the idea that differences between adult and child language learners may stem in part from initially minor differences that cascade throughout other aspects of language learning, we explored to what extent training adults on a previously unheard (novel) phonetic contrast results in improved learning of words that incorporate that contrast. Adults were assigned to either a TRAINED or CONTROL condition and trained distributionally, as in Maye and Gerken (2001). Both conditions were exposed to a bimodal distribution of phonetic sounds defined by voice onset time, but differed on whether the modes of the distribution mapped onto an existing phonetic contrast (the CONTROL condition: /d/ and /t^h/) or a novel contrast (the TRAINED condition: /g/ and /k/). We found that training on the phonetic contrast improved the learning of words beginning with that contrast, as well as the ability to discriminate a related contrast. These results have implications for how phonological abilities in adults affect other aspects of language learning, and for understanding how well distributional training enables phonetic generalization.

One interesting aspect of our findings is that as tasks became increasingly far removed from the original training, the ability to generalize diminished. The TRAINED group was able to generalize their phonetic learning to be able to discriminate a related but untrained contrast on a phonetic perception task, but when word learning was involved, they were only capable of learning words that began with the contrast they had been trained on. The CONTROL group was able to learn the *trained* and *untrained* contrast on the basis of the phonetic testing regime, but the resulting knowledge was more fragile than in the TRAINED group: their were unable to apply this ability to the problem of word learning. These results, in combination with the findings of other training studies (e.g., Bradlow et al., 1999; Maye & Gerken, 2001; McCandliss et al., 2002; Hayes-Harb, 2007), suggest that the ability to generalize phonetic learning (either to a related contrast or to another task) may depend strongly on the depth and nature of the training involved. It is possible that additional training would improve the ability to generalize even further. Relatedly, it is possible that our phoneme test did not measure phonetic category learning *per se*, and was more a measure of the raw ability to discriminate acoustically between two phonemes; if so, the limited generalization may have been due to the fact that our participants simply improved in their discrimination ability, but did not acquire phonetic categories in any reasonable sense (although the border between these two options is rather fuzzy). In general, the precise effect of training amount or type on generalization ability, and the nature of its dependence on the quantity and type of input, are matters for future study.

Our work was inspired in part by the idea that apparently major differences in language learning abilities may to some extent stem from smaller differences that have a cascading effect over time. While our findings are consistent with this notion, much work remains to be done to explore it more thoroughly, especially in the realm of adult language learning (research by Karmiloff-Smith and colleagues explores a similar idea in the area of language disorders). On one hand, it may appear unsurprising that being able to hear a phonetic contrast makes it easier to learn words that differ on that contrast. On the other hand, one might have expected phonological perception to have no effect on word learning: despite their poor perception, adults are arguably superior to children when it comes to acquiring vocabulary. Further work is essential, both for exploring whether linguistic abilities besides phonological perception affect other aspects of language, and for exploring whether phonological perception has effects on aspects of language besides word learning. This can include training studies like ours, as well as studies that evaluate how different aspects of language acquisition are affected by individual differences in adult phonetic perception (which are known to exist: see, e.g., McCandliss et al., 2002; Golestani & Zatorre, 2004; Perani, 2005; Golestani & Zatorre, 2009).

We conclude by noting an interesting puzzle: although the idea that deficiencies in one area of language acquisition can have cascading effects throughout other areas makes sense and is well-supported in the child acquisition literature (e.g., Tsao et al., 2004; Werker & Yeung, 2005; Fernald et al., 2006), so is the idea that jointly learning two aspects of language can improve performance in both (e.g., Feldman & Griffiths, 2009; Frank, Goodman, & Tenenbaum, 2009; Maurits, Perfors, & Navarro, 2009). However, the former implies that deficits in one area should propogate to another, while the latter implies that deficits in one area may be compensated for or overcome by skills or information from another. It is possible that both are true for different areas or in different ways, but as yet we know very little about the mechanisms or details underlying either, so it is difficult to know for sure. As usual, further research is necessary.

References

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Lang. Learning*, *56*(1), 9–49.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, *20*, 3–49.

Bradlow, A., Akahane-Yamada, R., Pisoni, D., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985.

Christiansen, M., Onnis, L., & Hockema, S. (2009). The secret is in the sound: From unsegmented speech to lexical categories. *Developmental Science*, *12*(3), 388–395.

Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*(12), 564–570.

Eimas, P., Siqueland, D., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Feldman, N., & Griffiths, T. (2009). Learning phonetic categories by learning a lexicon. In *31st Annual Conference of the Cognitive Science Society.*

Fernald, A., Perfors, A., & Marchman, V. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Dev. Psych.*, *42*(1), 98–116.

Fernald, A., & Simon, T. (1984). Expanded information contours in mothers' speech to newborns. *Dev. Psych.*, *20*, 104–113.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.

Frank, M., Goodman, N., & Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psych. Sci.*, *20*(5), 654–661.

Golestani, N., & Zatorre, R. (2004). Learning new sounds of speech: Reallocation of neural substrates. *Neuroimage*, *21*(2), 494–506.

Golestani, N., & Zatorre, R. (2009). Individual differences in the acquisition of second language phonology. *Brain & Language*, *109*, 55–67.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition*, *29*, 503–511.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research*, *23*(1), 65–94.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*(5), 219–224.

Hudson Kam, C., & Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tokura, Y., Kettermann, A., et al. (2003). A perceptual interference account of acquisition difficulties with non-native phonemes. *Cognition*, *87*, B47–B57.

Jamieson, D., & Morosan, D. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, *43*(1), 88–96.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398.

Kuhl, P. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.

Lively, S., Logan, J., & Pisoni, D. (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 Acoustic Society of America*, *94*, 21242–1255.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.

Maurits, L., Perfors, A., & Navarro, D. (2009). Joint acquisition of word order and word reference. In *31st Annual Conference of the Cognitive Science Society.*

Mayberry, R. (1993). First-language acquisition after childhood differs from second-language acquisition: The case of American Sign Language. *Journal of Speech and Hearing Research*, *36*, 1258–1270.

Maye, J., & Gerken, L. (2001). Learning phonemes: How far can the input take us? In *25th annual conference of the BUCLD.*

Maye, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101-B111.

Maye, J., Weiss, D., & Aslin, R. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134.

McCandliss, B., Fiez, J., Protopapas, A., Conway, M., & McClelland, J. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(2), 89–108.

McClaskey, C., Pisoni, D., & Carrell, T. (1983). Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics*(34), 323–330.

Mirman, D., Magnuson, J., Graf Estes, K., & Dixon, J. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*, 271–280.

Newport, E. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences*, *10*, 147–172.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical-access: Exemplar-based versus abstract lexical entries. *Psych. Sci.*, *12*, 445–449.

Perani, D. (2005). The neural basis of language talent in bilinguals. *Trends in Cognitive Science*, *9*(5), 211–213.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*, 111–123.

Shea, C., & Curtin, S. (2005). Learning allophones from the input. In *29th annual conference of the BUCLD.*

Singleton, J., & Newport, E. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, *49*, 370–407.

Snow, C. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of language* (pp. 257–276). Mahwah, NJ: Lawrence Erlbaum Associates.

Tan, L. (2003). Neural systems of second language reading are shaped by native language. *Human Brain Mapping*, *18*, 158–166.

Tsao, F., Liu, H., & Kuhl, P. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Wang, Y., Spence, M., Jongman, A., & Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, *106*, 3649–3658.

Weber, A., & Cutler, A. (2003). Lexical competition in non-native spoken word recognition. *JML*, *50*, 1–25.

Werker, J., & Lalonde, C. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Dev. Psych.*, *24*(5), 672–683.

Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*, 49–63.

Werker, J., & Yeung, H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, *9*(11), 519–527.