| Paper | Topic | Paradigm | Design | | Comments |
| --- | --- | --- | --- | --- | --- |
| Wade (2022) | Linguistic Convergence | Word elicitation prompted by a written hint and spoken hint. | 3 phases of elicitation of vowel /ai/.  Phase 1: Written hint and participant says word (prior production).  Phase 2: Spoken hint either in Midland or Southern accent (between participants). Participant says word.  Phase 3: same as phase 1 to investigate whether convergence is maintained  Participant vowel productions are measured for the realisation of the offglide along front diagonal of vowel space.  Analysis divided listeners’ by their region of origin (Southern or non-Southern)  The group who heard the hints delivered in a Southern voice shifted their /ai/ productions to be more Southern-like.  Interaction between listener origin, experiment phase and talkder dialect– Southern listeners shifted more than non-southerners. In phase 3 of Southern dialect condition, both Southeners and non-southeners show some but not full return to their baseline productions. | | Although this study is about production it goes to show that listeners maintain speaker models in their representation.  Listeners did not get any evidence of the southern /ai/ but both Southern and non-Southern listeners shifted their /ai/ productions towards what they believe were the talker’s origins.  This provides some support for the argument that listeners select from speaker models held in memory and behave accordingly.  If this is so, applying to our data, it would mean that listeners match the exposure talker to a best-fit model in their representation and hold quite strongly to it. That seems to me like holding on quite strongly to prior beliefs about the selected talker model and not adapting within the time frame of the experiment in spite of the evidence. |
| Escudero & Williams (2014) | L2 learning of vowel contrast | L2 Distributional Learning of vowels. Longitudinal study | (Pre-test, exposure, post-test) x2 with 6-month interval. Third (final) test without any further exposure, 12 months after first pre-test. 3 groups: Group1 = bimodal distribution over narrow variance; Group2 bimodal distribution over wide variance, Group3 exposure only to music.  Test tokens are natural stimuli from various talkers, with F1, F2 and *durational* cues. Training tokens were synthesised, duration held constant.  Post-test compared to pre-test in first 2 sessions to assess ST learning. Pre-tests compared between sessions to compare LT learning.  Performance was assessed on number of correct responses (XAB task).  While they found some evidence of LT training effects, it was only for the narrow bimodal group and not for the enhanced. This raises questions because they started from the position that wide variance training was superior to narrow variance training. | | This study set up a pre-test solely for comparison of accuracy between pre- and post-exposure. The difference with our approach is there isn’t any evaluation of where learners’ beliefs about the distribution lay and how this changed after each exposure. In my reading, the study says very little about participants’ prior beliefs because there isn’t an estimation of their categorization boundary along the relavant cue dimensions. Their concern was with improvement in categorisation. We are systematically tracking the shifts in perceptions as a consequence of their hypothesised learning of the phonetic distributions of the talker. In this study the test tokens are not plotted against training tokens to relate how the location of the distribution during exposure affects test. |
| Escudero et al., (2011)  N.B.  Both R2 and R3 cited this paper | L2 learning (vowel contrast) | L2 Distributional learning over short exposure. | Pre-test, exposure, post-test. 3 groups: Group1 = bimodal distribution over narrow variance; Group2 = bimodal distribution over wide variance, Group3 = exposure only to music. Post-test performance on XAB is compared with pre-test performance on XAB. | | My comments on this paper are the same as Escudero & Williams (2014) since they are essentially the same design but with different time frames.  I think a key difference between our testing of DL hypothesis and these L2 DL studies is in the specificity of the predictions. In these L2 studies they don’t seem to strictly adhere to the statistical learning principle. If learners are tracking statistics then the statistics of the cues before, during, and after exposure should matter and be measured. They hold to a general idea that exposure to wide or narrow distributions (I presume located within reasonable range) is what affects post-perceptual learning. So it doesn’t seem to matter so much where the test tokens are located relative to the exposure tokens.  A deeper analyses of their results would be interesting though. |
| Kraljic & Samuel (2007) | Perceptual recalibration | Lexically guided perceptual learning of /s/ and /sh/ | Listeners exposed to 2 talkers either with shifted /d/ or shifted /t/ in lexical contexts. Tested on d-t continuum with voices from both exposure talkers. Same design in experiment 2 but with shifted /s/ or shifted /sh/. | | R2 noted this paper for the number of trials (10) required for perceptual learning.  I think R2 wants us to note that other studies have shown the small number of trials for adaptation. They say “several studies on the amount of exposure needed for PL to take place are not cited”. We do discuss this and cite more recent work like Cummings & Theodore, which I believe is more specified than those before it. Besides that. I think this paper is testing a different question from us. |
| Poellmann et al., 2011 | Perceptual recalibration | Lexically guided perceptual learning of /s/ and /f/. Visual world paradigm | Investigates when recalibration emerges. Also asks if learning is gradual or step-wise.  The experiment exposes participants to shifted s or f while tracking their eye movements. Stimuli were delivered in blocks, with each block containing typical LGPL exposure trials (ambiguous s/f followed by clear s/f), test trials, less typical LGPL exposure trials (s/f embedded in temporary min pairs like gis/ftig – gis/fter ) and fillers (12 trials per block). There were 20 blocks in total.  Perceptual change is assessed by estimating an **average learning function over 2** blocks for each group (s-biased or f-biased). This is derived by subtracting the distance of fixations to the target from the distance of fixations to the competitor in each block and averaging them over 2 blocks  They found that in the first half both groups preferred looking at the s-item indicating that the stimuli was more s-sounding. The s-biased group always performed well by correctly looking at the target even in the earliest blocks. The f-biased group showed a sudden change in learning during block 11-12 and block 17-18 where they had a net preference in looks to the target. In most trials the f-biased group had a net preference for the competitor. | | This is an IcPhs proceedings paper which wasn’t published elsewhere.  This design is quite different from typical LGPL in that it has a test trial almost immediately after an exposure trial so that they are able to track the changes in perception trial-by-trial.  **I think this study is interesting and relevant to us because it aimed to track block-to-block changes in learning.** Their results did not provide strong evidence of rapid learning (they were trying to replicate the 10 exposure trials reported in Kraljic & Samuel 2007). While they report that perceptual learning is step-wise rather than gradual, I think their results can be challenged. I think this is why it didn’t appear in a standard peer-reviewed journal. Again, I think this study’s shortcoming is that they **do not locate the distribution of their stimuli nor relate the cue measurements of the exposure stimuli to the test stimuli** (or to the expectations of listeners before hearing the stimuli). This was the case with most LGPL studies in the earlier days and it was probably not until Drouin et al 2016 that they began to do that.  We could cite this study as one early attempt to detect the gradation of perceptual learning. |
| Witteman et al., 2013 | Accent adaptation | Cross-modal priming of German accented Dutch | Dutch listeners with greater or lesser experience with German accents respond to primes of strongly, medium, and weakly-accented primes.  Experiments tested listeners’ ability to generalise learning an accent from 1 talker to another. | | This study discusses **how prior experience/exposure affects how quickly one adapts to accented speech**.  They note the speaker-dependent nature of ST adaptation because naïve listeners exposed to a German accented talker did not give much advantage in a later test phase which had a different talker.  However, this conclusion contradicts the finding that experienced listeners showed stronger priming.  I think this study is distantly related to our paper because of its research questions. R2 mentioned this paper as an e.g. of studies that looked into prior knowledge, which they criticise us for not acknowledging.  That’s similar but we are estimating prior perception and tracking the direction of change due to recent exposure. This paper recognises the role prior experience plays and divides its participants into groups by experience. It is not measuring the degree or direction of adaptation. We can cite/acknowledge that people have thought of accounting for prior experience before but they don’t usually try to estimate the representation in terms of cue distributions. |
| Zhang & Holt 2018  (experiment 1) | Adaptation to manipulated VOT-F0 correlations.  How do changing statistical relationships between cues affect perception. | Holt’s dimension based stat. learning | Listeners given 3 blocks of stimuli where VOT either correlated with F0 in the expected direction (canonical) or in the reverse direction. F0 ranges were manipulated to be either relatively higher (high F0 condition) or relatively lower (low F0 condition).  Prior to exposure listeners in both the high F0 and low F0 group were subjected to a pre-test to check how they categorised the VOT-F0 continuua. 120 total trials in pre-test  The 2 groups shared one set of VOT stimuli coupled with a mid-level F0 value (210 Hz). This mid-level F0 value is treated as the low F0 (coupled with short VOTs) for the high F0 condition and the high F0 (coupled with long VOTs) value in the low F0 condition.  Test stimuli was placed in one VOT location –10ms coupled with either mid-level F0 or high F0 in the high condition or with mid-level and low F0 for the low condition. | | Categorisation of the ambiguous stimuli showed that the respective groups adapted their perception relative to the F0 range that they were exposed to.  Crucially, in the ambiguous test stimuli, mid-level F0 was categorised more as /b/ in the high F0 group and more as /p/ in the low F0 group. This suggests that listeners normalised the F0 to reflect the exposure distribution. I thought this was the more interesting part of the experiment.  The pre-testing was done to determine how F0 affected the categorisation of stimuli in the absence of correlational statistics. Listeners in respective groups were given two uniform VOT continuaa paired with the 2 F0 levels. This could be taken as their prior expectations of the stimuli.  The results show that listeners can change the way they categorise ambiguous stimuli according to correlations of recent exposure.  Because there isn’t much distributional manipulation other than the correlation of VOT and F0, this study is quite different to ours. They do have tests mixed up in each exposure block but they do not test for intermittent changes.  They do not test predictions of DL. |
| Logan et al., 1991 |  |  |  | |  |
| Eisner & McQueen 2006 |  |  |  | |  |
| Reviewer 3 | | | | | |
| Anderson 1982 |  |  | |  |  |
| Gauthier et al., 2007 |  |  | |  |  |
| Goudbeek et al., 2008 |  |  | |  |  |
| Goudbeek et al., 2009 |  |  | |  |  |
| Guenther & Gjaja 1996 |  |  | |  |  |
| Heathcote et al., 2000 |  |  | |  |  |
| Logan 1988 |  |  | |  |  |
| Maye & Gerken, 2000 |  |  | |  |  |
| Maye et al, 2003 |  |  | |  |  |
| McMurray et al., 2009 |  |  | |  |  |
| McMurray et al., 2009b |  |  | |  |  |
| Newell & Rosenbloom 1981 |  |  | |  |  |
| Rescorla 1988 |  |  | |  |  |
| Toscano & McMurray 2010 |  |  | |  |  |