Appel à projets incitatifs « au fil de l'eau »

Titre du projet: Perceptual assimilation in second-language word learning and the contribution of visual input to the assimilation process

Porteur(s) du projet et Laboratoire(s): Chotiga Pattamadilok (LPL), Michael Tyler (Western Sydney University)

1) Objectives, hypotheses and methods

Infants are born with the potential to learn to discriminate between any pair of phones (i.e., consonants or vowels), and they gradually learn to tune in to those that signal a meaning difference in the native language, and tune out from those that do not. The vestiges of this attunement can be observed in adults when they are presented with contrasting phones in a language that they have never encountered before. Japanese native speakers have notorious difficulty perceiving the difference between English /r/ and /l/ (Goto, 1971) and English native speakers with the Hindi dental-retroflex contrast /d̪/-/ɖ/ (Werker & Logan, 1985). Most European French speakers seem to have difficulty with the English interdental fricatives /θ/ (as in "think") and /ð/ (as in "this"). Observation of French-accented English would suggest that most perceive /θ/ as a version of their French /s/ category, and /ð/ as French /z/, although some individuals may perceive /θ/ as /f/ and /ð/ as /v/. However, there is surprisingly little empirical data on how European French speakers categorize English /θ/ and /ð/ and discriminate them from other English consonants.

There are a number of theories of how the native language shapes the perception of non-native phones, and the success of acquiring those phones as new categories when learning a new second language (L2). Here we concentrate on the Perceptual Assimilation Model (Best, 1995; Best et al., 2001; Tyler et al., 2014), and its extension to L2 learning, PAM-L2 (Best & Tyler, 2007). According to PAM, perception is tuned, over development, to the phonetic and phonological properties of the native language. A non-native phone is either assimilated to a native-language category as a good, acceptable, or poor exemplar (*categorised*), heard as speech but as no one native category in particular (*uncategorised*), or not heard as speech at all (*non-assimilable*; e.g., English listeners' perception of non-native click consonants; Best et al., 1988). Pairs of phonologically contrasting non-native phones form assimilation types based on how each one is assimilated to the native language. A *two-category* assimilation is formed when each non-native phone is assimilated to a different native category, and an *uncategorised-categorised* assimilation when one is categorised and the other is not. When both are assimilated to the same native category, then the contrast is a *category-goodness* assimilation when one is perceived a poorer example of the native category than the other, or a *single-category* assimilation when they are perceived as equally good or equally poor examples of the category. PAM predicts the following for discrimination accuracy: Two category > uncategorised-categorised ≥ category goodness > single category. These predictions have been consistently confirmed in studies in support of PAM (e.g., Best et al., 2001; Tyler et al., 2014).

PAM-L2 predicts the likelihood of new L2 category formation using these assimilation types as a basis. Category development is most likely for contrasts involving an uncategorized phone and for category-goodness assimilations. When the learner does not manage to create a new L2 category for one of the phones in a single-category or category-goodness assimilation, minimal pairs should end up being homophonous in the L2 lexicon. For example, a French speaker who does not distinguish /s/ and /θ/ may hear the same word for *thing* and *sing*. However, some studies have shown that homophony may be *asymmetrical* (Weber & Cutler, 2004), that is, one member of the minimal pair could be correctly recognized faster than the other, although the origin of such asymmetrical homophony is still under debate (Escudero et al., 2008).

Although PAM and PAM-L2 consider speech perception to be amodal, studies in support of PAM have focused almost exclusively on the auditory modality. Yet, speech perception is not solely an auditory phenomenon. When available, input from other modalities is also used. In literate individuals, speech is tightly associated with two main sources of visual information, that is, speaker's articulatory gestures and orthography. Both kinds of visual input have been reported to influence the way speech is perceived. As illustrated in the McGurk effect (McGurk & MacDonald, 1976), an illusory percept (e.g., /da/) can be created by the fusion of the auditory stimulus (/ba/) and the facial cues (a face articulating /ga/), thus suggesting an integration of the auditory and visual information during speech perception. The fact that orthographic knowledge can also influence the way speech is processed has been reported in several studies. For instance, Seidenberg and Tanenhaus (1979) showed that participants were faster to recognize two spoken words as rhymes when they were spelled similarly (e.g., tie-pie) relative to dissimilar spellings (tie-rye). Similar findings have been reported in many other metaphonological and speech recognition tasks (Pattamadilok, Morais, Colin, & Kolinsky 2014; Tyler & Burnham, 2006; Ziegler & Ferrand, 1998). These observations have been argued to reflect either an online activation of the orthographic representation during speech processing or a structural change of the phonological system induced by reading acquisition (Taft, 2011).

More related to our issue is the fact that the contribution of visual information significantly increases in difficult listening conditions, i.e., when auditory information alone could not lead to a satisfying outcome (Pattamadilok et al., 2011; Ross et al., 2007). As described above, learning to discriminate non-native phonemes could be considered as an ambiguous speech processing situation. **The aim of the proposed study is to investigate perceptual assimilation phenomenon in second-language word learning and examine how visual information from speakers' articulatory gestures and word spelling contributes to improving the learning outcome.**

The first experiment we will consist of perceptual assimilation and discrimination tests on contrasts involving English /θ/ and /ð/. This will serve two purposes: 1) to establish perceptual assimilation patterns for these consonants (using both French and English labels) and to test PAM predictions for discrimination of these consonants against confusable English consonants; 2) to provide pre-test data for Experiment 2 – i.e., participants who assimilate /θ/-/s/ as single category, category goodness, or two category (using English labels), will be invited to return for the second experiment. We expect there to be substantial interindividual variability – most participants will assimilate the English fricatives to /s/ and /z/, others will assimilate them to /f/ and /v/ or /t/ and /d/, and some will have acquired the fricatives as a new L2 categories. We hypothesise that the discrimination accuracy will vary as a function of the participant's individual assimilation type (see Tyler et al., 2014): two category > category goodness > single category.

The aim of Experiment 2 is to test how L2 word learning is affected by perceptual assimilation (as identified in Exp 1), orthography, and presentation modality. The experiment will be based on the novel word learning paradigm of Escudero et al. (2008), where participants learn labels for novel objects, followed by a visual-world eye-tracking task to assess the time course of recognition. We will invite participants from Experiment 1 for whom the /θ/-/s/ contrast is: 1) single category; 2) category goodness, or; 3) two category (i.e., they have acquired a new L2 category for /θ/). For single-category participants, we hypothesise that looks to target versus looks to competitor will be the same for words contrasting /θ/ and /s/, regardless of which one is the target and which is the competitor. That is, overlapping portions of words containing /θ/ and /s/ will be symmetrically homophonous. For category-goodness participants, looks to target will be higher than looks to competitor when the target contains /s/, but they will be similar to each other when the target contains /θ/ - i.e., they will show asymmetrical homophony. For two-category participants, looks to target will be higher than looks to competitor regardless of which consonant is the target. Homophony should be reduced when visual information (orthography or articulatory gesture) is provided, relative to auditory-only presentation.

Methods

Experiment 1: Perceptual assimilation and discrimination of English fricatives by French native listeners.

Participants. We will continue to collect data until we have sufficient numbers of participants for Experiment 2 (24 for each assimilation pattern: single category, category goodness, two category) We anticipate around 150 participants from the LPL participant panel, which will allow us to examine individual differences in perceptual assimilation.

Stimuli and apparatus. Auditory stimuli will be consonant-/a/ syllables, recorded by Michael Tyler, a native speaker of Australian English. The critical stimuli will contain the consonants /θ/, /s/, /f/, /ð/, /z/, and /v/, with an additional six English consonants included as filler items, and three syllable tokens per consonant category. Stimulus presentation will be controlled by a computer running jsPsych, with an external sound card and high-quality headphones.

Procedure. Participants will complete six AXB discrimination tasks, followed by two categorisation with goodness rating tasks, one using French labels and another using English labels. On each trial in an AXB discrimination task, the participants are presented with three syllables. The consonant in the first and third syllable are always different, and the participant's task is to match the consonant in the middle syllable with the consonant in the first or the third syllable. They will complete the task for six contrasts: /θ/-/s/, /θ/-/f/, /θ/-/t/, /ð/-/z/, /ð/-/v/, /ð/-/d/. In the categorisation task, the participants hear a single syllable and select a category label from a set of forced choices. E.g., on hearing /sa/, they would select "S" from a list of French consonant labels. The syllable is then played a second time, and participants rate the goodness-of-fit of the token to the category that they chose on a scale of 1 (poor) to 7 (excellent). The categorisation will then be repeated for using English labels. The entire session, with time for breaks, will take around 90 minutes. This procedure has been successfully used to study perceptual assimilation in studies spanning 30 years (e.g., Best et al., 1988, 2001; Tyler et al., 2014).

Experiment 2: The role of perceptual assimilation on second-language word learning and the contribution of visual cues

Participants. There will be 72 participants in total, 24 in each of three groups: 1) single category; 2) category goodness, and; 3) two category. Each participant will undergo three training conditions 1) auditory only; 2) auditory+orthography, and; 3) auditory+articulatory movements. The order of the training conditions will be counterbalanced across participants.

Stimuli and Apparatus. The 60 novel objects (20 for each condition) will be taken from the novel object and unusual noun database (Horst & Hout, 2016: http://www.sussex.ac.uk/wordlab/noun). Following Escudero et al. (2008), in each condition there will be five novel words for each of four consonant categories: /θ/, /s/, /t/, and /k/ (e.g., *thindesh, sinrold, tinchen, kinwig)*. /θ/ and /s/ were chosen as potentially confusable pairs, with /k/ and /t/ as distractors. The novel words will be recorded in audio-visual format by the same speaker as for Experiment 1.

Procedure. The procedure will follow that of Escudero et al. Participants will first compete a word learning task, consisting of 10 blocks of 60 trials each, where a novel word is spoken and participants click on the corresponding picture. On each trial participants will be given feedback on whether they were correct or incorrect and then the novel word will be repeated with the correct picture. The novel word will be presented according to condition: auditory only, auditory+orthography, or auditory+articulatory. Figure 1 illustrates the learning procedure.

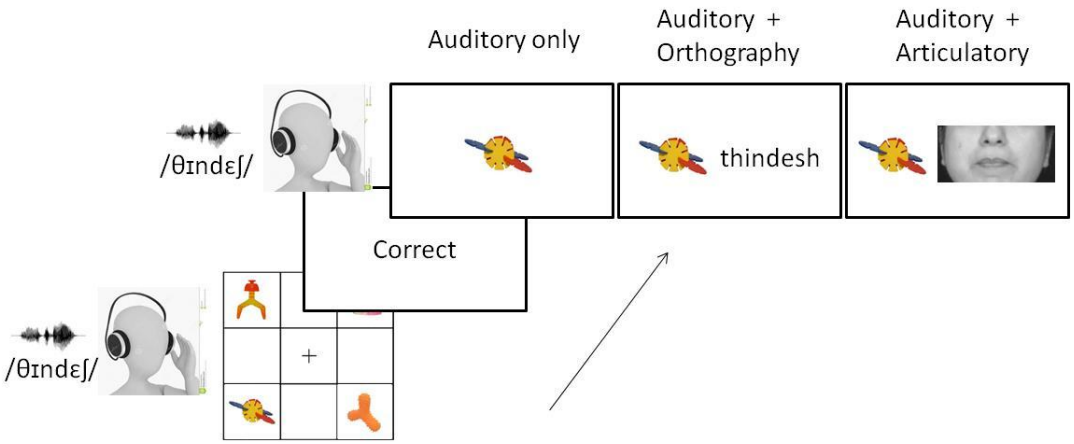

Figure 1: The learning procedure

In the test phase, on each trial participants will see a picture in each of four quadrants of the screen and an auditory only instruction to click one of the pictures (e.g., click on the *thindesh*). Their eye movements will be recorded using a Tobii T120 eye tracker and responses collected using E-Prime. No feedback will be provided. Participants will complete each condition on a different day. Each session will take around 1 hour.

For both Experiments, a control group consisting of native speakers of English will be tested in Australia by Michael Tyler using his own funding.

2) Originality of the project

This project is the start of a new collaboration between Chotiga Pattamadilok (LPL) and Michael Tyler (Western Sydney University), who is visiting LPL during his sabbatical. It is the first study to investigate how multimodal perception, orthography, and perceptual assimilation might influence the acquisition of new lexical items in a second language, and it has great potential to provide new insights into the interface between phonology, the lexicon and the contribution of visual cues. A second novel aspect of the study is that, surprisingly, there has been no previous systematic investigation into the perception of English interdental fricatives by continental French speakers. A large-scale study examining the breadth of individual differences in perceptual assimilation will clearly establish how these phonemes are perceived by native continental French speakers.

3) Complementarity of expertise (if researchers from different BLRI/ILCB labs are involved; this is encouraged but not required)

Michael Tyler is an author of studies testing the Perceptual Assimilation Model, and he is co-author of the influential PAM-L2 model (Best & Tyler, 2007; 806 Google Scholar citations). He has studied spoken word recognition and lexically guided perceptual retuning, but his main expertise is in the perception of consonants and vowels. As such, he approaches the interface between phonology and the lexicon from the bottom up. His expertise is complemented by Chotiga Pattamadilok, whose research on the role of orthography in speech perception approaches the same questions from the top down.

4) Relevance of the project for the BLRI/ILCB as well as the potential impact of the expected results

If successful, the results of the novel word learning experiment will be of interest to editors of the high-impact journals that published the studies on which it is based (Journal of Memory and Language, Journal of Phonetics). The results also have potential application for foreign language learning pedagogy. For example, students could be screened prior to learning to assess whether their

perceptual assimilation pattern is conducive to the introduction of orthographic materials, or whether targeted perceptual training is required first. Furthermore, the question of how English interdental fricatives are assimilated by French native speakers, and the consequence of that perceptual assimilation on discrimination, is likely to attract interest from the scientific community and from the general public.

5) Budget and co-funding

Detailed demands

- Compensation for participation (Exp 1: 150 subjects * 1.5h * 10€/h = 2250€; Exp2: 24 subjects * 1hr/session * 3 sessions * 10€/h = 720€. Total = **2970€**)
- Financial support for a research assistant who will be in charge of subject recruitment and data collection (2350€ brut/month * 2 months = **4700€.**The study will require at least 300 hrs of data collection)
-Technical support for experiment setup and data analysis (the eye tracking part).

Co-funding: When he returns to Australia, M. Tyler will run control conditions with native speakers of English using his own funding.

References

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*(2), 775-794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 345-360.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r.". *Neuropsychologia, 9*(3), 317-323.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*, 2711-2724.

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393-1409.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746–748.

Pattamadilok, C., Morais, J., Colin, C., & Kolinsky, R. (2014). Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain and language, 137*, 103-111.

Pattamadilok, C., Morais, J., & Kolinsky, R. (2011). Naming in Noise: The Contribution of Orthographic Knowledge to Speech Repetition. *Frontiers in Psychology, 2*, doi: 10.3389/fpsyg.2011.00361. doi:10.3389/fpsyg.2011.00361

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. (2007). Do you see what I am saying? Exploring visual enhancement of speech in noisy environments. *Cerebral Cortex, 17*, 1147–1153

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5*(6), 546-554.

Taft, M. (2011). Orthographic influences when processing spoken pseudowords: theoretical implications. Frontiers in Psychology, 2, doi: 10.3389/fpsyg.2011.00140

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica, 71*(1), 4-21.

Tyler, M. D., & Burnham, D. K. (2006). Orthographic influences on phoneme deletion response times. Quarterly Journal of Experimental Psychology, 59(11), 2010-2031.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1-25.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37*, 35-44.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review, 5*, 683-689.