On the Limits of L1 Influence on Non-L1 Listening: Evidence from Japanese Perception of Korean

Anne Cutler [1,2], Jeesun Kim [2] and Takashi Otake [1,3]

[1] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [2] MARCS Auditory Laboratories, University of Western Sydney, NSW, Australia; [3] E-Listening Laboratory, Tokorozawa, Japan
anne.cutler@mpi.nl, j.kim@uws.edu.au, otake@e-listeninglab.com

Abstract

Language-specific procedures which are efficient for listening to the L1 may be applied to non-native spoken input, often to the detriment of successful listening. However, such misapplications of L1-based listening do not always happen. We propose, based on the results from two experiments in which Japanese listeners detected target sequences in spoken Korean, that an L1 procedure is only triggered if requisite L1 features are present in the input.

1. Introduction

Listeners process spoken language in ways which are subtly tailored to suit the phonological structure of their native language. This has been best documented with respect to the segmentation of continuous speech into its component words. This draws on rhythmic structure, and because rhythm differs across languages, speech segmentation procedures are likewise language-specific.

Thus syllabically based segmentation of speech was demonstrated for French listeners but not for English listeners (Cutler, Mehler, Norris and Segui, 1986); the latter were shown to use stress-based segmentation instead (Cutler and Norris, 1988; Cutler and Butterfield, 1992; Cutler, Mehler, Norris and Segui, 1992). Since French and English had long been held up as the classic examples of “syllable-timed” and “stress-timed” rhythm respectively (Abercrombie, 1967), this pattern suggested that the segmentation differences could be based on rhythmic differences. This in turn predicted that a different rhythmic structure could also be the basis of a language segmentation procedure. Indeed, the moraic rhythm of Japanese proved to be accompanied by mora-based segmentation by Japanese listeners (Otake, Hatano, Cutler and Mehler, 1993; Cutler and Otake, 1994; Otake, Hatano and Yoneyama, 1996).

The discovery that segmentation procedures are language-specific provided a potential explanation for the frequent experience that segmenting speech in a non-native language (henceforth: non-L1) is hard. Where the rhythmic structure of the non-L1 differs from that of the native language (L1), use of the L1 processing strategy will be counter-productive. Experiments confirmed that non-L1 listeners did not necessarily segment speech in the same way as L1 listeners. The French speech materials which produced syllabic segmentation from French listeners did not do so with English (Cutler et al., 1986), Japanese (Otake et al., 1996) or Dutch listeners (Cutler, 1997). French listeners presented with Japanese did not segment moraically (Otake et al., 1993), and neither did English listeners (Otake et al., 1993; Cutler & Otake, 1994). The way English listeners responded to English materials was not replicated by French (Cutler et al. 1986) or Japanese listeners (Cutler and Otake, 1994).

Instead, listeners tended to apply their L1 listening strategies inappropriately to the non-L1 input. French use of syllabic procedures with English (Cutler et al., 1986) and Japanese (Otake et al., 1993) was parallelled by Japanese application of moraic segmentation to French, Spanish, and English (Cutler and Otake, 1994; Otake et al., 1996). Following those studies, there have been many demonstrations of inappropriate application of L1 segmentation to non-L1 input, even when the non-L1 is a high-proficiency L2 (Weber and Cutler, 2006).

The experiments revealed not only that listeners from languages with different rhythmic structure produced non-native-like response patterns, but also that listeners from (related or unrelated) languages with similar rhythmic structure produced L1-like response patterns. Note that the establishment of similarities in perceptual reflections of rhythmically based segmentation across otherwise unrelated languages offers a possible route to test the Rhythmic Class Hypothesis, which holds that languages fall into a limited number of groups defined by rhythmic similarity. Attempts to provide a phonetic measure for testing this hypothesis (e.g., Low, Grabe and Nolan, 2000; Ramus, Nespor and Mehler, 1999) are labour-intensive and hence difficult to apply to large samples, but can be rather unreliable with small samples; an alternative testing approach is thus very useful.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

The first such test involving completely unrelated languages exploited the close resemblance between the (quite simple) phonological structure of Japanese and that of the languages of the Dravidian family – Telugu, Tamil, Malayalam, etc. The phonological similarity between these unrelated language families has been remarked upon for over a century (Caldwell, 1856; Shiba, 1973), and was sufficient to motivate the prediction that listeners from Dravidian languages would show evidence of the moraic segmentation previously observed only in Japanese listening. Cross-linguistic experiments in Telugu and Japanese by Cutler, Murty and Otake (2003; Murty, Otake and Cutler, in press) indeed revealed significant similarity in segmentation behavior across the two listener populations.

A parallel series of studies is underway with French and Korean. Recall that French listeners use a syllable-based segmentation strategy. Syllabic segmentation has also appeared in experiments in Korean (e.g., Yoon and Derwing 1995, 2001). The phonological structure of French and Korean differ; French allows complex onsets and a variety of coda types, and has a mid- to large-size phoneme inventory with many vowels. Korean also has a mid-size vowel inventory; but it allows no onset clusters, few coda clusters, and also constrains which consonants can be codas. Onsets may be consonants or null, medials are vowels or diphthongs, codas one or two consonants or null. Thus a syllable can be (C1) (S) V (C3), where parentheses signal optionality, C1 is any consonant other than /ŋ/, S is one of the semi-vowels (/w, j/), and C3 is (an unreleased variant) of any of /p/, /t/, /k/, or one of the three nasals /m/, /n/, /ŋ/ or the liquid /l/ or /s/ only before a word or suffix beginning /s, s´/ (Kim and Davis, 2006). Different as these phonologies are, Yoon and Derwing's results suggest that syllabic segmentation is as effective in Korean as it has been shown to be in French.

In the present investigation we extend the cross-linguistic comparisons by presenting Korean materials from our forthcoming studies to Japanese listeners. If Korean indeed proves to pattern like French, with syllabic segmentation the procedure of choice for L1 listeners, then Korean and Japanese will mismatch in the same way as French and Japanese did in the experiments of Otake et al. (1993, 1996). It is still not necessarily the case, however, that Japanese listeners will then apply their segmentation procedures to Korean in exactly the same manner as they were shown to do for French. The phonological differences allow for many other factors to exercise an effect on the listening results.

The segmentation research program made use in large part of the fragment detection task (Frauenfelder and Kearns, 1996), in which response times and miss rates are recorded as listeners detect prespecified targets – word fragments – in spoken input. In the present study we use this task too. It has the great advantage that it can be performed without any knowledge of the input language, so that testing of non-L1 listeners is feasible. Targets may be presented visually or auditorily with equivalent results (Otake et al., 1993), and the task permits both RT and miss rate measures. With non-L1 listeners, RTs are sometimes very long or miss rates are high so that the latter measure may be more informative (Otake et al., 1993; Cutler et al., 2003).

A moraic segmentation pattern would reveal itself as selective difficulty for CVC targets in CV-initial words, as found in the Japanese and Telugu experiments; CVC targets in CVC-initial words, and CV targets in either word type, should be easier, and equivalently so. The classic syllabic effect (e.g., in French) is a crossover: the conditions showing a match between target type and word structure (CVC targets in CVC-initial words or CV targets in CV-initial words) are easier than mismatched conditions (CV in CVC-initial words, CVC in CV-initial words). All other patterns of response fail to address the predictions from the rhythmic hypotheses (e.g., simple effects of target type or word type are not predicted).

2. Experiment 1

2.1. Materials and design

Eighteen Korean words, comprising nine pairs, were selected as experimental stimuli. The words within each pair were chosen such that they all began with the same three segments (e.g., *pem*) but differed in that one word had CVC/CV structure (e.g. *pem/cwu*), while the other had CV/CV structure (e.g., *pe/ma*). A further 250 filler words were chosen. From these items a total of 71 sequences were constructed: seven practice sequences, and two experimental sets (A and B) of 32 sequences varying in length from two to six words. Each set of 32 included 14 filler sequences without occurrence of the specified target or with a target early or last in the sequence; the remaining 18 sequences contained one of the chosen stimulus words, in 2nd, 3rd, 4th or 5th position.

All practice and experimental sequences were recorded to Digital Audio Tape in a sound-attenuated cubicle by a female native speaker of Korean. All sequences were preceded by a spoken target syllable. There were two different target syllable types: CV (e.g., *pe*) or CVC (e.g., *pem*). Each experimental stimulus word occurred twice, once in Set A and once in Set B. Sets A and B differed in the target assignment for the experimental words. If *pema* was assigned target *pem-* in Set A, it was assigned target *pe-* in Set B. The target assignments for its pair would be the reverse (*pe-* in A, *pem-* in B). The targets for the practice and filler sequences were also CV or CVC syllables.

2.2. Procedure

The subjects were tested in separate sound-attenuating carrels in a quiet room, either individually or in pairs.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

They were told that they would hear a list of foreign words following a sound sequence specifying the auditory target. They were instructed to listen for a word beginning with the spoken target, and then to respond as fast as possible by pressing the response key provided.

The materials were presented over high-quality headphones from a Sony TCD-D10 DAT player which was interfaced with a personal computer running NESU experiment control software. Timing pulses were recorded on the second channel of the tape aligned with the onset of each experimental target word; these pulses were inaudible to the participants. The computer's clock was started by each timing pulse and stopped by a participant's key-press. A timeout was registered if the participant did not respond within 1.5 seconds.

2.3. Participants

Forty student volunteers from Dokkyo University, Soka, Japan, took part. All were native speakers of Japanese and none had learned Korean prior to the experiment. They received course credits for their participation. Twenty heard the item sets in AB and 20 in BA order.

2.4. Results and Discussion

Mean response times and mean missed responses per condition are shown in Figures 1 and 2. Analyses of variance were conducted across participants and across items for each type of measure. Although the pattern in each figure suggest that overall the listeners found CVC targets easier to detect than CV targets, the main effect of target type was not significant and in fact the analyses revealed no significant effects at all for either RTs or errors. The results thus reveal neither a moraic nor a syllabic response pattern in this experiment.

We had chosen to present the targets only auditorily, given that the earlier studies of Otake et al. (1993, 1996) had found no difference in the pattern of results with visual and auditory specifications. The visual targets in those cases were in the Roman alphabet appropriate for the English, French and Spanish spoken input, and not in Japanese orthography; this avoided any possibility of L1 orthography artificially inducing use of the L1 strategy. The Roman alphabet is familiar enough to Japanese students from advertising use, or from language courses. However, that option was not available to us in the present case. For Korean materials the Roman alphabet is not appropriate, and the listeners cannot read Korean orthography. After the experiment, informal debriefing of the participants revealed that most of the auditorily specified targets had been perceived as bimoraic. Only the CVC targets, however, were intended as bimoraic; the CV targets should have been perceived as mono-moraic, but were apparently heard as CVV instead. The results we observed could therefore have arisen because the CVC targets constituted a better match to the CVC-initial words, and were thus easier than the CV targets.

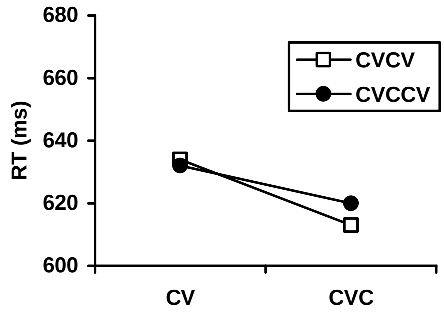

Figure 1: Experiment 1: Mean RTs as a function of word structure (CVCV, e.g., *pema*; CVCCV, e.g., *pemcwu*) and target size (CV e.g., *pe-*; CVC e.g,. *pem-*).

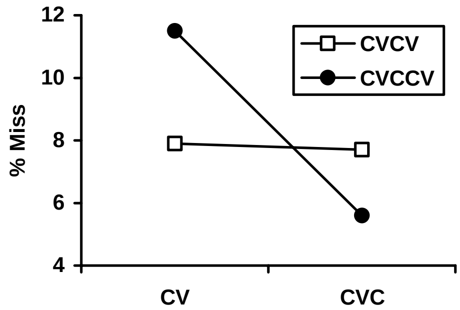

Figure 2: Experiment 1: Mean proportion of missed responses as a function of word structure (CVCV, e.g., *pema*; CVCCV, e.g., *pemcwu*) and target size (CV e.g., *pe-*; CVC e.g,. *pem-*).

Auditorily specified targets are spoken separately, not extracted from recordings of longer spoken words. Such extraction produces unnatural-sounding tokens, and also turns the task into a simple acoustic matching response. This is not the required processing level; to show effects of L1 phonology, the task needs to be carried out at a phonological, not an acoustic processing level. Matching of a separately spoken syllable to the same syllable in a word meets this criterion.

Any naturally spoken syllable is longer in isolation than within longer words. However, the CV syllables had been intended by the speaker as CV; they should not have been long enough for CVV. Measurements were carried out of the targets and first syllables of items. These revealed that all target syllables were longer than the same syllables in the words, but the length ratio was just under 2:1 for CVC-initial items, but in every case above 2:1 for CV-initial items (the average duration of a CV target was 259 ms, and the ratio of target to the same syllables in the target-bearing words varied from 2.2:1 to 3.9:1). We therefore decided to attempt to achieve the perceptual effect intended for these targets by repeating the experiment with the target specifications compressed to equivalent ratios in the CV and CVC cases.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

3. Experiment 2

3.1. Materials and Procedure

These were as in Experiment 1 except that, using the algorithms available in the CoolEdit 2000 software package, all the auditory target specifications were compressed (preserving pitch and amplitude contours of the original tokens). We used a compression factor of 1.8 for the CV targets (reducing the targets to 55.55% of their original duration), and .999 for the CVC targets (leaving the duration virtually untouched but ensuring that any acoustic consequences of a compression having been applied were equivalent for CV and CVC targets).

3.2. Participants

A further 36 student volunteers from the same population, meeting the same criteria as for Experiment 1, participated in return for course credits. None had taken part in the preceding experiment. Eighteen heard the item sets in AB and 18 in BA order.

3.3. Results and Discussion

Mean response times and the mean number of missing responses per condition in Experiment 2 are shown in Figures 3 and 4. It can immediately be seen that the pattern is quite different from that in Experiment 1. Analyses of variance showed that in RTs, the only effect significant across both subjects and items was the interaction of word structure with target type (F1 [1,34] = 4.06, p = .05; F2 [1,16] = 4.83, p < .05); t-tests on the components of this interaction revealed no significant target type effect for CV-initial words, but a significant advantage for CV over CVC targets for CVC-initial words (t [35] = 2.12, p < .05). Note that this is not the interaction predicted for a moraic effect, which would have resulted in a difference for CV-initial but not for CVC-initial words.

Thus the compression manipulation was in part successful in that it enabled the listeners to recognise an acoustic match with the CV-initial words. But not only the CV item results changed – those for CVC did too.

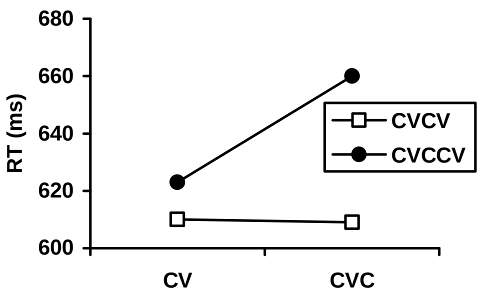

Figure 3: Experiment 2: Mean RTs as a function of word structure (CVCV, e.g., *pema*; CVCCV, e.g., *pemcwu*) and target size (CV e.g., *pe*-; CVC e.g,. *pem*-).

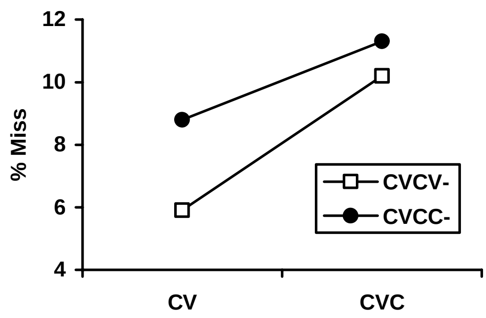

Figure 4: Experiment 2: Mean proportion of missed responses as a function of word structure (CVCV, e.g., *pema*; CVCCV, e.g., *pemcwu*) and target size (CV e.g., *pe*-; CVC e.g,. *pem*-).

3.4. Combined Analysis

Finally, the results of Experiments 1 and 2 were jointly analysed. There were no significant differences in the overall difficulty of the two experiments (grand mean RTs differed by less than one ms across the two responses sets, and mean miss rates by less than 1%), but the impression given by Figures 1 through 4 of a different pattern of responses in the two experiments was confirmed by significant interactions of the variable Experiment with the other independent variables. In the RT analyses the interaction between Experiment and Word Structure was significant across participants (F1 [1,72] = 5.18, p < .03; F2 [1, 16] = 3.39, p < .085), as was the interaction between Experiment and Target Type (F1 [1,72] = 4.72, p < .04, F2 [1,16] = 2.65, n.s.). In the miss rate analyses only the interaction between Experiment and Target Type was significant, but this effect was robust across participants and items (F1 [1,72] = 6.81, p < .02, F2 [1,16] = 10.59, p < .005). Sub-analyses confirmed the strongest interaction to be a difference across experiments in the target effect in CVC-initial words (RTs: F1 [1,72] = 4.38, p < .04, F2 [1,8] = 2.96, n.s.; Miss rates: F1 [1,72] = 6.93, p <.02, F2 [1,8] = 12.85, p < .005).

4. General Discussion

Our two experiments have shown that Japanese listeners can detect targets in Korean words, but show no tendency to apply their L1 segmentation procedures to the non-L1 speech input when doing so.

In the first experiment there was a weak tendency only for one effect: detection of CVC targets seemed somewhat easier than detection of CV targets. The CV targets as naturally recorded by the Korean speaker were perceived as possibly bimoraic (CVV) by the Japanese listeners, making them mismatch the CV-initial words, in which the first syllables were very much shorter. In the miss rate analysis the highest miss rate was observed for CV targets in CV-initial words, emphasising the fact that there was in this case a perceived mismatch.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after full paper review

In the second experiment we attempted to remedy this apparent mismatch by adjusting the timing of the target specifications. The results again showed an unexpected target effect, but in the reverse direction: CV targets were now in general rather easier than CVC.

Note that the alteration in pattern of results across the two experiments has itself been very informative. The compression of the overlong CV targets to make them a closer acoustic match to the initial syllables of the CV-initial items was the only change in the materials. The experiments were otherwise identical; the running order and word sequences were all unchanged, and the CVC targets, except for having been run through the compression software, were also unchanged. Yet the pattern of results proved to be quite different.

This suggests, first of all, that response patterns are strongly affected by the overall probabilities in an experiment. In Experiment 1 listeners found the CVC targets comparatively the easiest to deal with. Even though in Experiment 2 the CVC targets were as before, the CV targets were then easier. These results join the many demonstrations of alterations in responses as a function of changing probabilities in the stimuli. Tasks such as fragment detection are highly susceptible to such strategic effects, since they involve a simple response and are not very engaging of participants' attention. If there are simple cues which listeners can latch onto to help them perform the required detection, these will be used. In phoneme detection, closely related to fragment detection, effects of lexical status of the target-bearing items can be made to come and go by altering the fillers in the experiment; /d/ in *dip* and *dap* is detected equally fast if all the words and nonwords in the experiment are monosyllabic, but /d/ is detected faster in *dip* than in *dap* if the fillers include some bisyllabic words (Cutler, Mehler, Norris and Segui, 1987). More variety engages listeners' attention to the items to a greater extent.

We note here that an earlier version of the current experiments comprising 152 trials (10 practice trials and two sets of 71 experimental trials) was run with Korean, Japanese and English-speaking listeners. This study, about twice as long as the ones reported here and more than twice as long as the fragment detection experiments run in French etc., proved too long for all listeners. This too is a cautionary lesson about the structure of the whole experiment in phonological listening studies.

However, the principal result of the present study is that L1 influence has not been observed. In neither experiment did a moraic effect appear. The listeners did their best to find a match to the targets, but in doing so they did not apply their L1 moraic strategy. The only way in which our results might be said to conform to predictions from rhythmic hypotheses was that the Japanese listeners showed no trace of a syllabic effect. The effects that did show up were effects of target size, which are most likely to be strategic in nature.

Why did Japanese listeners, who have previously been shown to apply their L1 segmentation to input in various forms of non-L1, fail to segment the Korean words in their preferred L1 manner? In fact the literature on segmentation shows other cases of similar listener restraint. In publications, for natural reasons, positive results enjoy more attention than such negative findings; but the negative results do exist. For instance, Murty et al. (in press) report that Telugu materials which included phonological patterns impossible in Japanese – such as coda clusters – elicited no moraic segmentation

Again, whereas the Japanese listeners in Cutler and Otake's (1994) study detected both consonants and vowels faster in moraic rather than non-moraic position in Japanese, in English the Japanese listeners only detected consonant targets faster in "moraic" position (/n/ in *candy* rather than *canopy*) – with English vowels (/o/ in *kiosk* rather than *abolish*) this did not happen. In the vowel sequences the Japanese listeners did not hear successive vocalic elements (which might have triggered a moraic response), but apparently a vowel-glide-vowel.

That is, in Cutler and Otake's (1994) study the listeners did not perceive the input as allowing a moraic analysis. We propose that the explanation of the present results likewise lies in the degree to which the listeners could analyse the input as conforming to requirements for application of a moraic segmentation procedure. The phonemic sequences in the Korean input were putatively too far from Japanese L1 phonological legality for a match to L1 expectations to be possible.

We assume, more generally, that listeners will only draw on L1 listening heuristics if encouraged to do so by the structure of the speech input they are presented with. L1 experience has encouraged listeners to develop these heuristics because speech recognition thereby becomes more efficient. Where the phonological structure of a non-L1 input affords a match to structural expectations from the L1, then, the same experience suggests that listening to this non-L1 should benefit from application of the familiar heuristics. This may prove in fact to be the case; but if the apparent match is in fact a spurious one, interference will result instead of benefit. In either case, it is the presence in the non-L1 input of specific features, perceived by the listener to be L1-like, which has triggered the use of L1 listening procedures.

The practical import of the findings of this kind of L1 transfer is of course not so much in listening to any non-L1 but in listening to a deliberately acquired L2. If interference from L1 persists in L2 listening – as we know it does – then this can have serious consequences for the listener's communication goals. We speculate that it may be possible for the status of the features in the L2 which trigger L1 influence to alter, and indeed to alter in either direction. That is, features which are initially assimilated to L1 categories may with increasing experience come to be perceived as distinct from the L1

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review

categories (so that the L1 procedure no longer comes into play), but also L2 features not initially perceived as similar to the L1 may turn out to share regularities of phonological patterning with L1 features (and thereby come to trigger the relevant L1 procedures after all).

When listeners cannot accommodate non-L1 speech input to L1 structure, they may of course be better able to hear the real acoustic structure. In the case of phonetic segments it is most obviously possible to apply L1 categories, since all languages have vowels and consonants; yet here too, as Best, McRoberts and Sithole (1988) have shown, segments which cannot possibly be incorporated in L1 category structure may be accorded acoustically veridical processing. Rhythm is a relatively abstract level of structure; the results to date have shown many cases in which L1 rhythmic processing is applied to non-L1 input, but these results do not allow us to isolate phonological features which trigger application of L1 rhythmic procedures. It is possible that we may actually come to learn more about these trigger features, which are likely to be quite precisely specified, by exploring instances, such as the present case, in which application of L1 procedures does not occur because of putative mismatch between L1 and non-L1 structure.

We propose that for both theoretical and practical reasons it is now important to establish the range of features, or the minimal structural elements, which, if encountered in a non-L1, will satisfy the criteria for application of each rhythmic segmentation procedure.

5. Acknowledgements

This research was supported by the NWO SPINOZA project “Native and Non-native Listening”, JSPS grant (C) 11610566, and ARC grant DP0453143. We thank Natasha Warner, Chris Davis, and Maarten Jansonius for assistance with planning and conduct of the study, and Taehong Cho for phonetic consultations.

6. References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

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

Caldwell, R. (1856, reprinted 1976). *A Comparative Grammar of the Dravidian or South-Indian Family of Languages.* London/Madras: University of Madras Press.

Cutler, A. (1997). The syllable's role in the segmentation of stress languages. *Language and Cognitive Processes,* 12, 839-845.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language,* 31, 218-236.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology,* 19, 141-177.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24, 381-410.

Cutler, A., Murty, L., & Otake, T. (2003). Rhythmic similarity effects in non-native listening? *Proceedings of the Fifteenth International Congress of Phonetic Sciences*, (vol. 1, pp. 329-332). Barcelona.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance,* 14, 113-121.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824-844.

Frauenfelder, U.H., & Kearns, R.K. (1996). Sequence monitoring. *Language and Cognitive Processes,* 11, 665-673.

Kim, J., & Davis, C. (2006). Literacy acquisition in Korean Hangul. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of Orthography and Literacy* (pp. 377-387). Hillsdale NJ: Erlbaum.

Kim, J., Davis, C., & Cutler, A. (forthcoming). Perceptual tests of rhythmic similarity. II: Syllable rhythm.

Murty, L., Otake, T., & Cutler, A. (in press). Perceptual tests of rhythmic similarity. I: Mora rhythm. *Language and Speech*.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language,* 32, 358-378.

Otake, T., Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese listeners. In T. Otake & A. Cutler (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies* (pp. 183-201). Berlin: Mouton de Gruyter.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.

Shiba, S. (1973). Daravida-go to Nihongo (Dravidian languages and Japanese). *Jinbun Ronsaa,* 22.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America*, 119, 597-607.

Yoon, Y.B., & Derwing, B.L. (1995). Syllable saliency in the perception of Korean words. *Proceedings of the XIIIth International Congress of Phonetic Sciences* (vol 2, pp. 602-605). Stockholm: KTH & Stockholm University.

Yoon, Y.B., & Derwing, B.L. (2001). A language without a rhyme: Syllable structure experiments in Korean. *Canadian Journal of Linguistics,* 46, 187-237.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after full paper review