

WORKING MEMORY, PHONEMIC CODING ABILITY AND
FOREIGN LANGUAGE APTITUDE: POTENTIAL FOR
CONSTRUCTION OF SPECIFIC LANGUAGE APTITUDE
TESTS – THE CASE OF CANTONESE

Erica Chan

Chinese University, Hong Kong

Peter Skehan

Chinese University, Hong Kong

Gwendolyn Gong¹

Chinese University, Hong Kong²

Abstract:

A re-awakening of interest in foreign language aptitude in recent years has also seen some reanalyses of the construct. Prominent in these reanalyses is revised conception of memory and aptitude, and particularly, the role of working memory. A number of studies have explored the relevance of non-word repetition tests as measures of this type of memory. Building

Ilha do Desterro	Florianópolis	nº 60	p. 045- 073	jan/jun 2011
------------------	---------------	-------	-------------	--------------

on this work, the present article reports on the development of language-specific non-word repetition tests. It is proposed that non-word repetition tests based on the phonological structure of the language to be learned will be particularly effective at predicting second language learning. Non-wordness, in this view, is defined with reference to the target language. The article reports on the development of such tests, focussing on the case of Cantonese, a tonal language, where approaches to the characterisation of non-words are particularly challenging. The article reports on methods of overcoming these difficulties. It also argues that such an analysis of working memory brings this construct close to Carroll's characterisation of one of the components of foreign language aptitude – phonemic coding ability.

Introduction

Many years ago Cronbach (1975), writing within mainstream psychology, lamented the separation of the discipline into two camps: researchers who focused on learner differences, and researchers who were more concerned with common processes and universals. He pointed out that these two groups rarely talked to one another, to their mutual disadvantage. Within applied linguistics the study of aptitude has provided a more recent example of this very same phenomenon – on the one hand, there is an individual differences perspective (Skehan 1989; Dörnyei, 2005), but on the other, there are acquisitional processes (reviewed in Mitchell & Myles, 2004). One approach emphasizes differences while the other strives to uncover general processes. The irony, of course, is that they both may be dealing with the same thing.

The conventional approach to aptitude has been to administer large numbers of potential predictor tests, correlate them with one another, and also with achievement/proficiency tests, and then attempt to establish the constructs which underlie the aptitude

tests and most effectively predict language learning success. It was on this basis that the most celebrated of aptitude researchers, J.B. Carroll, proposed his four factor view of aptitude (Carroll, 1965) as consisting of phonemic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory. Almost fifty years after this account of aptitude first appeared, it is still the most influential portrayal of the area available. It also has, at its heart, the paradox that the clarity of the identification of the four aptitude constructs is not matched in a clear connection between each separate component and the different sub-tests which can be found in aptitude test batteries. As a result, the sub-field of aptitude study has been seen as contributing more to differential psychology, the first of the areas mentioned by Cronbach, since the principal value of aptitude tests has been prediction, rather than explanation. The “theory” of aptitude has been inferred from the account Carroll provided. Although one can make links between it and psychological processes, this is not facilitated by the hybrid nature of most actual aptitude tests, which sacrifice “purity” of measurement in order to boost the level of prediction.

In any case, at the time that Carroll was researching, conceptions of acquisition, first and second, were vastly different from now. Since then the field of second language acquisition has grown enormously, and our understanding of the psycholinguistic processes implicated in acquisition and learning have deepened considerably (De Bot, Lowie, & Verspoor, 2006). It was with these things in mind that an earlier account (Skehan, 2002) proposed that it would be timely to approach aptitude testing more theoretically, and driven by putative second language acquisition processes, rather than by interpretations of factor analyses of aptitude test batteries. In that respect, the following stages might be of relevance:

- Input processing
- Noticing
- Pattern identification
- Extension
- Complexification/ Restructuring
- Integration
- Error avoidance
- Automatisatation
- Repertoire and salience creation
- Lexicalisation

The first of these emphasizes the preliminary stage of input processing, an area investigated in detail recently by VanPatten (1996). The next five, from Noticing to Integration, are concerned with structural and interlanguage development (Klein, 1986). The final four are concerned with the capacity to use language, in real time, with a general emphasis on proceduralisation, culminating in a capacity to access a dual-mode system (Pawle & Syder, 1983; Skehan, 1998). In this view, the “chunks” which underlie language may be analyzable into syntactic patterns, but they may not be accessed as such during real-time performance. Instead, they are drawn on as formulaic wholes so that speed of processing is enhanced, and computational demands are reduced (Schmitt & Carter, 2004).

The purpose of exploring the above developmental processes is to follow a simple claim: if there are individual differences in any of the above areas, then, starting from a processing perspective, we have

a *prima facie* case for considering that there may be an associated component of foreign language aptitude. The proposal that we start from these processes also suggests that it would be useful to attempt to link existing aptitude test components to these processes, since this might reveal (a) which areas are already covered, however imperfectly, and (b) which areas need further work. However, this earlier sketch was vague, and the purpose of the present paper is to explore, in just one area, how it can be made, if not precise, at least less vague!

Characterising Interlanguage Processes: The centrality of working memory

As a preface to this section, it is useful to recall the analysis of comprehension provided by Clark and Clark (1977). They suggest that (first language) listeners take in phonological representations of speech, analyse this material, extract meaning, and then purge working memory of the “slice” of speech that has been processed. In other words, input is transformed into meaning, and then the decks are cleared, so to speak, because the next slice of input is about to arrive. This raises a fundamental issue for language processing: how do listeners, let alone acquirers, cope with the real-time problem, and, in the case of acquirers, who may want to extract useful things from the input to cause change in underlying interlanguage systems, i.e. to focus on form (Doughty, 2001), how can they make such extractions under the time pressure they are under?

Clearly the first stage to consider is that of input processing. Focussing on spoken language, a number of issues become apparent at this stage. First, there is the very considerable problem of input segmentation. Learners have to find ways of analyzing the stream of sound into smaller units which can then be operated upon more

effectively. This is vital since units which have been segmented may then trigger comprehension processing strategies (Clark & Clark, 1977; VanPatten, 1996), parsing, and lexical access. Each of these subsequent processes will be important to “deliver” some sort of message to the listener. And the existence at this first stage of strategic processing raises an issue which will recur: deterministic, exhaustive, syntactic-based processing may be avoidable if other methods of meaning-extraction are sufficient. In other words, immediately we have the prospect that a focus-of-form (Long & Robinson, 1998) may not be implicated in natural, ongoing comprehension. Faster processes may be given “first chance” to operate and if they are able to construct likely intended meaning, the available, more powerful syntactic processes may be unused, or only partially used.

It will be assumed here that the next vital stage in acquisition is that of noticing (Schmidt 1990, 2001). The previous analysis of the early stage of input segmentation and preliminary analysis suggests that the language learner, especially, (even beyond the native speaker) will be under some pressure, and so, following Sachs’ (1967) early work, will tend to discard exact form very quickly in order to have working memory resources available for the next input. As a result, Schmidt’s point is that noticing is not the default processing option. Something additional has to happen for noticing to occur, and further, that if noticing does not occur, there will be no stimulus for the interlanguage system to change. Hence his concern that the learner extracts some feature of the input as worthy of further attention, and possibly retains this feature for additional work.

Here it is useful to distinguish between two types of noticing which might occur. At a first level, there might be what could be termed “fuzzy” noticing, where something is noticed as being of interest but without analytic progress. Something salient may be

noticed, for example, but not understood, e.g. Schmidt's (Schmidt & Frota, 1986) vague realization that pronoun forms in Brazilian Portuguese seemed to be varying strangely. Alternatively, there may be a more completely realized noticing, in which a specific feature of input is extracted, and may be itself the basis for further analysis, e.g. a precise noticing of the subjunctive form. Perhaps as a final point, it is important to link this discussion with the work of Merrill Swain (1995), who has argued for the importance of noticing in output. Schmidt tends to focus on input, while Swain is more concerned with the memory traces which are laid down when learners are speaking, and have to confront, through the messages they want to formulate, gaps in their interlanguage. They may, as a result, notice what they don't know, and, as a result, do something later about this gap in their knowledge.

The reason for exploring these early processing stages in detail is because of the connections they potentially offer for aptitude. In that respect, one of the key developments in foreign language aptitude has been the greater realization we now have of the functioning of working memory (Harrington and Sawyer 1992; Miyake & Friedman 1999; Robinson, 2002). Earlier work, e.g. Skehan (1982), used a passive short-term memory construct, and digit span tests. These did not correlate with language learning success. The situation has changed considerably since then as more structured and active conceptualizations of *working* memory have become influential (Miyake & Shah, 1999; Baddeley, 2007). Several features of this development are relevant.

First, although there is some disagreement amongst working memory theorists about model details, most agree, following Baddeley (Baddeley, 1986; Gathercole & Baddeley, 1992), that there is structure to working memory, and that it contains a processing

component and also storage buffers, in at least two modalities (verbal and visual), with another component, the episodic buffer, concerned with links with long-term memory (Baddeley, 2007). Second, there are resource limitations (but see Robinson 2003, for a different view on this), such that the amount of material that can be held in working memory buffers is limited, and also that the amount of computation possible for the processing component (as well as the span of immediate consciousness) is also limited. The material in working memory may be available to immediate consciousness, (indeed, may *be* immediate consciousness), and it may be operated upon very fast, but there cannot be very much of it. In this, it contrasts with the more capacious, but less accessible material in long-term memory. Third, part of the role for the processing component is to *direct* working memory resources, including its own operation (!), to ensure that the resources available are used effectively. In this there is an acceptance that there will be a tradeoff between (amount of) attention directed to processing, and the amount of material that can be held in working memory buffers (Baddeley, 2007).

Obviously the key issue is to consider what the role of working memory is within language processing. Miyake and Friedman (1998) make the claim that, with its processes of analysis and immediate storage, it is central. We can return to the immediately preceding analyses of input processing and noticing to demonstrate how this is so. Regarding input processing, while one can see that the process of segmentation will partly be a bottom-up process, it is also the case that top-down processing, when possible, will be advantageous. For this to happen more effectively, it may be useful to have larger stretches of language available so that more extended patterns can be accessed in order to cope with ambiguous input strings. Regarding lexical access, it may well be that whether working memory means

activated records within LTM or not, it is still useful to have more working memory resources, and faster processing to access LTM addresses to enable the lexical accessing to take place which can then trigger syntactic analyses. In other words, at some detail one can see how there are advantages for second language input processing if working memory resources are more extensive and more powerful.

As indicated earlier, one of the central problems in second language acquisition is that communication (and therefore more communicative approaches to language pedagogy) will emphasise meaning, and there is no natural reason, during ongoing communication, to pay attention to form. In other words, if in comprehension meaning can be extracted, or, in production, through communication strategies or formulaic language, composed, there is little purpose in trying to notice aspects of form. Influences on noticing therefore become particularly important, as the discussions in Schmidt's work make clear. In an ideal world, one would like learners, while communicating, to simultaneously attend to form and meaning and to explore mappings between the two. Researchers therefore have explored how such a focus on form can be facilitated (Doughty, 2001). In this respect, it is clear, once again, that larger working memory will make noticing more likely to occur. One way in which this is so is that if working memory resources are more effective, the primacy of meaning can be handled and there may be attentional resources left over for a focus on form (VanPatten, 2003). This would be a form of "surplus resources" interpretation. It would be unselective, and simply available. This could be either because central executive resources within working memory are faster, or it could be because there is simply greater buffer capacity, or both. But the consequence might be that buffer-encoded material could remain active for longer, thus making it more likely that input

features could be extracted *for form*. As Doughty's (2001) analysis makes clear, negotiation for meaning and recasting presuppose some capacity for working memory to maintain activity for material which is the focus for the feedback which is being provided. In other words, the enhanced potential for noticing that is involved has to assume a memory system which operates in a way to enable the necessary form-meaning mappings to be made.

But there is also the slightly separate issue of what one does with the material held in working memory. In this respect, the key issue is that working memory is not simply a workspace to facilitate ongoing processing: it is also the gateway to long term memory. This is crucial. To notice something is simply part of the fleeting nature of life. To notice and to make connection with existing structures in long term memory is to offer the prospect of changing the contents of long term memory and to achieve interlanguage development. Without such connections, the prospect of ongoing communication and the insights it might potentially provide leading to underlying change disappear. Working memory, in other words, is the necessary component to enable current noticing to trigger change and subsequent development. And, once again, it is clear that more effective working memory operation will make the formation of such connections more likely.

In a way, the discussion so far has somewhat avoided mention of aptitude. Now, though, it is time to return to this core area. We noted above that previous approaches to aptitude were test-battery driven, using statistical techniques to winnow down the potential tests available to the subset which made the greatest predictive contributions to language learning success (and which therefore tacitly define the structure and components of aptitude). Here, with the analysis of input processing and noticing, we see a totally different approach to

identifying aptitude components. It is simply to say that one starts by working with potential acquisitional processes and moves to identifying areas of cognitive functioning which might be implicated. Then comes the crucial step. If an area of cognitive functioning is important *and there are individual differences in that area* then potentially we have an aptitude component where differences between learners in the cognitive capacity concerned can connect with differences in level of second language learning success. The justification for regarding working memory as a component of aptitude therefore rests on its functioning within acquisitional processes, coupled with the probability that learners will vary in the component in question. With working memory, although there have been many changes in conceptualization over the years, there is a general (though not universal) consensus that individual differences do exist, and that Miller's original proposal of "the magic number 7, plus or minus 2" may be questioned in respect of detail, but not in general insight: some people have more effective working memories than others. What is interesting is the possibility that the meaning of "effective" is more complex now, since speed of processing, executive control, and workspace size (and structure) all combine to influence what level of effectiveness is involved for any one individual. Hence the different approaches to, for example, reading span tests as different ways of characterizing this combination of working memory structures and operations. We will return to this fundamental insight several times: are there relevant cognitive abilities for the processes of second language acquisition? And, relatedly, are there individual differences which are relevant to this ability?

Phonemic Coding Ability

Carroll's four factor theory of aptitude gives considerable importance to phonemic coding ability. Originally, this factor was

called sound-symbol association ability, and this is the label that is used in Pimsleur's LAB (1966), but examination of factor analytic results led Carroll (1965) to change the description of the construct to the one we know today – the ability to analyse sound so that it can be retained for more than a few seconds. The construct description is a remarkable insight, but it is not entirely borne out when one examines actual test implementation. This involves sub-tests such as Phonemic Script learning and Spelling Clues (Carroll & Sapon, 1957). The first requires a system for transcription to be learned, while the latter provides curiously spelled versions of words, e.g. *ernst*, and requires the identification of the 'hidden' word, 'earnest'. The former seems to test sound-symbol association ability, pure and simple, while the latter draws upon general L1 vocabulary, together with the ability to relate verbal to auditory material. Another test, number learning, does require memory and is also proposed as a measure of auditory alertness. In sum, the promise of the construct-based insight, an ability to code unfamiliar material so that it can be retained for more than a few seconds, is not totally fulfilled. Coding is interpreted as sound-symbol association only – a rather limited characterization, but one which has been shown subsequently to connect with dyslexia (Sparks & Ganschow, 1993), itself an example of such symbol-linked coding problems.

What are not developed in the actual test procedures are any method of handling coding or analysis of material without symbols, and any notion of how 'unfamiliar' might be operationalised. So there is here, as in other areas of the MLAT, the problem that clarity of constructs is not matched by clarity or at least purity of measurement. This is a pity, since the construct description emphasises ways in which fleeting memories can be retained. This seems close to working memory, and especially the use of the episodic buffer as a means of

holding on to working memory contents for a greater length of time. In addition, working memory researchers have linked this construct to language learning. For example, Gathercole and Baddeley (1993, p. 74) propose that “the long-term learning of the sound structures of novel, phonologically unfamiliar words depends on the availability of adequate representations of the sound patterns in the phonological loop.” This directly implicates the phonological loop component, and relates this to the capacity to handle unfamiliarity. Indeed, Gathercole (1995) relates degrees of non-wordness to ease of repetition and learning. Carroll (1965) seemed to anticipate developments in cognitive psychology many years ahead. All the more disappointing then that MLAT sub-tests did not really capitalise upon this insight, inevitably providing potential for the development of working memory/ phonemic coding tests which are more informed by more recent research findings.

Measuring Working Memory, and the Importance of Specific Languages

As indicated earlier, several components have been proposed for working memory. There is reasonable consensus that there is a central executive, modality-linked buffer systems, and an episodic buffer concerned with long-term memory connections. There are also somewhat divergent approaches to measurement in each of these areas. The first, the central executive, is typically measured with span tasks, in which immediate memory is involved, but so is computation of some sort (e.g. judging the truth value of sentences, carrying out some mathematical calculation). The score is the number of items that can be recalled even though there is active central executive involvement. Phonological memory, i.e. the most relevant modality-based measure, is typically measured by non-word repetition tests

where items which conform to the phonological structure of a language, but which do not actually exist are used, such as 'tablus', or 'acklar'. This procedure has been the focus for considerable research (see Gathercole 2006, for review), and many salient variables have been identified, e.g. the syllable length of the non-words, the word-likeness of the non-words, or the frequency of the syllables which make up the word.

The emphasis in the earlier discussion of phonemic coding ability was the retention of verbal material. It may be that central executive influences are relevant, but the main focus is on handling verbal material. For that reason, we will focus here on the development of a test of the phonological buffer. As will be argued below, this connects most clearly with phonemic coding ability. It is the area where there is potential for progress in attempting to characterize unfamiliarity more precisely. But there is a complication which we need to address at the outset, and this relates to one of the long-term issues with aptitude tests. This concerns specificity. Existing aptitude tests (e.g. MLAT (Carroll & Sapon, 1959), DLAB (Petersen & Al-Haik, 1976), CANAL-S (Grigorenko, Sternberg, & Ehrman, 2000) are general purpose instruments. They are proposed as usable for the learning of any language -- in passing it should be noted that these tests, in base form, generally assume L1 English learners, but translations of some of these tests are available. In this way, an L1 English learner of German could take these tests, as could an L1 English learner of Japanese. Then, since it has been a cornerstone of aptitude research since Carroll's original work that aptitude tests should not be used for selection, or rejection from a course, but rather to provide estimates of the length of time that would be needed for a particular individual to reach a certain criterion level of performance, the same base aptitude test would be used to predict different lengths of time to reach some

criterion in German (a language strongly related to English), and a comparable criterion in Japanese (a language which is more distant).

There are obvious practical reasons why using all-purpose tests of this sort is desirable. They enable one test to be developed, and then it is assumed that the same set of abilities is used for the different languages which might be learned. In other words, if the abilities associated with, e.g. inductive language learning, are relevant, then it is assumed that they will be relevant in the same way for German and Japanese, but they might have a more difficult task in the latter case, and might therefore predict that more time would be needed to reach a criterion. There is also the advantage that if one aptitude test is used, that it can be translated without undue complication if a different L1 will be involved. In any case, developing one aptitude test is difficult enough: developing a series of different tests would be excessively difficult.

But there is still the possibility that developing language-specific tests might have advantages, and could, ultimately, be worth the additional effort. Of course, this requires a justification of why specific tests connect in an advantageous manner with specific target languages. One needs to know, in other words, why tests of memory, or language analytic ability, or phonemic coding ability are improved if they incorporate something specific about the language being learned. This additional effort would be eased, though, if a means could be found to produce such tests following some consistent method where the methods to proceed for different language combinations are well-understood. It is in fact the purpose of this article to explore such a possibility with the development of tests of the phonological component of working memory. There may be comparable arguments that can be advanced for other components of foreign language aptitude. For now, we will only focus on this more modest and circumscribed aim.

Central to this possibility for the construction of a test of non-word repetition is the nature of a non-word. One of the most salient variables in working memory tests of this sort is that different non-words vary in how language-like they are perceived. In other words, not all non-words are equal, and performance on tests based on, for example, English-like non-words is higher (with L1 English participants) than performance on less English-like non-words.

This raises an interesting possibility. Nonwords based on English are different from nonwords based on other languages. Given the centrality of working memory as a component of foreign language aptitude, it is, of course, possible to develop phonological memory aptitude tests which are based on nonwords in the L1. But it is intriguing to consider what would happen if non-words were developed which are based on the language to be learned. At a practical level, such a framework for test construction might allow the generation of broadly comparable tests based on different languages quite easily. One would only need to have methods of generating and rating relatedness of non-wordness to the target language to be able to proceed. In a sense, this develops a proposal made by DeKeyser and Juffs (2005) where they argue that it may be useful to have domain specific working memory tests. In their case, the focus is on different aspects of language performance, e.g. noticing feedback, comprehension, production, but the point is general – all-purpose working memory tests may need to be supplemented or replaced by more focussed instruments which lead to more selective predictions.

But there is also a more important theoretical advantage. Any aptitude test is going to need to contain a component assessing auditory skills. However, even with early research it was clear that the capacity simply to make sound discriminations is not the central issue. Carroll (1965) argued cogently that it is not sound discrimination

per se which contributes to aptitude – it is the capacity to handle sound beyond individual segments, and above all, the capacity to analyse sound so that it can be retained. Carroll (1965), in other words, as argued earlier, was very close to the construct of working memory here. He also emphasised sound-symbol correspondences, but basically, the need to be able to retain sounds was the core to this aptitude construct. Now, with non-word repetition tasks, and phonological memory, we are dealing with essentially the same construct.

Which brings us to one more facet of Carroll's phonemic coding ability test – it concerns the ability to retain *unfamiliar* sound that was at the heart of the ability. This supports the proposal for the construction of a non-word repetition test based on the phonological rule system of the language to be learned. Essentially, such a test will incorporate the sound structure of the target language, and if this unfamiliar sound system is different from the L1, variation in the ability to deal with such sounds could be extremely important.

Constructing a Language-specific Non-word Repetition Test: the challenging case of Cantonese

The basic set of procedures implicit in what is being proposed is fairly simple:

- Identify the target language
- Clarify the phonological rules for word formation in that language
- Generate a range of candidate non-words in that language
- Possibly have these rated for their 'wordlikeness'

- Use these words for the production of a non-word repetition test

If these procedures are applied to English, they can proceed smoothly and are even assisted by the existence of a large number of non-words which are available from previous research, often accompanied by statistics regarding degree of English wordlikeness (Gathercole, 2006). In the present case, the languages which are implicated were English and Cantonese. For English, the above procedures and resources rendered the task straightforward. Interestingly, for Cantonese the problems are rather different, and it is the difficulties that were involved in solving those problems that is covered in this section.

For Cantonese speakers, ‘words’ are defined more by Chinese characters than by sounds (since one writing system is used to represent Cantonese and Mandarin, and a range of other dialects and languages). These characters represent entire syllables, not individual sounds. Following this portrayal of ‘words’, one cannot have one-syllable non-words, because to many Cantonese speakers, to speak of a word without a character is meaningless. One can, though, produce two-syllable nonwords, where the two syllables, each representable by a character, do not occur in combination, although they might occur singly, or in combinations with other syllables (i.e. Chinese characters). So the entire sequence of sounds which defines the nonword may be novel, but the component syllables are not, much as when, with English, one uses a nonword such as *bookdog*. Cantonese non-words produced in this way will be referred to as *syllable nonwords*. Originally nonword repetition was preferred to word repetition because non-words do not provide the same basis for the participants to make associations which would aid memory. Such behaviour would cloud the measurement of a pure phonological

capacity. The use of non-words, such as ‘bookdog’ obviously changes the way associations are made but does not prevent them. For that reason, it does not seem an ideal method in the present circumstances.

The obvious answer here is to proceed with the sound combinations that characterize Cantonese. But this brings us up with a different problem. Cantonese is a tonal language, and depending on the system one uses, contains six to nine tones. So to a non-Cantonese speaker, what appears to be the same word, though spoken with a different intonation, can mean several different things, each meaning associated with a different Chinese character and, to a Cantonese speaker, entirely different words (although to a beginner learner of Cantonese, these may appear to be exactly the same word). For example, *gwai* refers to “tortoise” if one pronounces it with a high-level tone, but the same consonant-vowel combination refers to “ghost” if it is pronounced with a high-rising tone. In passing, it should be said that there are methods of representing Cantonese words alphabetically, through systems which have been developed to capture tonal information, e.g. the Yale Romanization system. These, however, are completely unfamiliar to Cantonese speakers themselves.

So this raises the possibility of defining phonological Cantonese non-words in two different ways. First a non-word may be defined as a combination of sounds which seems to meet Cantonese phonological rules, but which is not used. Following Stokes et al. (2006), those non-words can be sub-classified as IN-syllable and OUT-syllable nonwords. IN syllables are those CV combinations that occur in the Cantonese syllabary. An example of IN syllable non-words would be *bap*, a completely non-existent sound sequence in Cantonese. *Ba* is an IN syllable because *bam*, *ban*, *bat* and *bak* all exist in Cantonese. There are also what Stokes, Wong, Fletcher, and Leonard (2006) call OUT syllable non-words with CV structures

that do not appear in the Cantonese syllabary. *We* is an OUT syllable because it is impossible to find any words with a CV structure as such, i.e. *we*, *wem*, *wep* and *wek* are all non-occurring. A non-word, *weng*, can be created in this way.

What started out as a problem in fact turns into an opportunity. We have learned from English-based research with non-word repetition tests of the phonological buffer that certain variables are salient, and increase or decrease difficulty. In a sense, this is useful information for the construction of verbal memory foreign language aptitude tests since one can use these findings to attempt to calibrate difficulty in the tests concerned. Most of these findings also apply, in principle, to Cantonese, e.g. the relevance of syllable length, the importance of ratings of wordlikeness. But Cantonese provides the additional possibility of constructing non-word repetition tests using additional variables. What is proposed here is that a non-word repetition task can be built up of separate sequences of:

- Syllable-based nonwords
- IN-syllable nonwords
- OUT-syllable nonwords

This enables a non-word repetition test of the form:

- Syllable based non-words
- Presentation of two syllable syllable-based non-words (starting with two items, and increasing to (say) five items, over two cycles)
- Presentation of three-syllable syllable based non-words (starting with two items and increasing to (say) four items, over two cycles)

- In-syllable non-words, following the above presentation schedule
- Out-syllable non-words, following the above presentation schedule

In this way, one would have, simultaneously, a comprehensive measure of nonword repetition as the basis for a predictor of second language learning of Cantonese, and one would also have an inherent research design, since one could explore the *relative* effectiveness of the different bases for the production of non-words. It will be interesting to see if constructing a non-word based on the target language, but using different characterizations of non-word leads to differentially effective prediction.

To clarify the claims being made here, it may be useful to give examples of each of the types of non-word. Syllable based non-words are of the following form (with tone information supplied numerically):

ping4sou3

baa6gei1

coi4fu4

muk6jin6

dou6ling5

faa3de1ngon6

The first example here ‘means’ ‘apple mathematics’, since it combines these two individual Chinese characters into a ‘word’ that none of the authors is familiar with! Tone representations are 1:high level; 2: mid rising; 3: middle level; 4: low falling; 5: low rising; and 6: low level, following the Sidney Lau (1977) system for representing tone in Cantonese. All these non-words received high Cantonese-like ratings, of greater than 3.5 on a 1-5 scale.

IN-syllable words are exemplified as follows:

gyu5tak1	pim5fap1	hem5paam1
leoi3kyu1faap6	noe3kaan1map6	nyu3te1toek6

In this case, the CV combination in each syllable does exist in the Cantonese syllabary even though the word itself does not exist in the language.

Finally, OUT-syllable words are exemplified as follows:

weng5myun1	jou5moeng1	woi5foi1
beoi3jot1hui6	woe3jou1myut6	foei3tei1lui6

In this case, the sound combinations in each syllable, although permissible through Cantonese phonology, do not actually exist, in any tone.

As a final point in this section, it should be noted that this ‘blueprint’ for the construction of a nonword repetition language aptitude test does exploit the structure of Cantonese phonology, and as a result the test that has been constructed is more extensive and also more systematic in its sampling, since three types of non-word are involved. Cantonese, in other words, provides more opportunities for non-word construction. Essentially, it is only the second and third categories of Cantonese non-word, the in-syllable and out-syllable words, which would be available for languages such as English (tonal languages, e.g. Mandarin, Vietnamese, might provide greater opportunities, like Cantonese). But the point here is that the method is general in scope. A difficult task has been addressed in the present case because in this way the method can be shown to be workable under unusual circumstances. Most target languages would be easier to work with.

Redefining the Construct of Phonemic Coding Ability

Having explored the particular case of developing a Cantonese non-word repetition test, we can return to the issue of how Carroll's (1965) original conception of phonemic coding memory has been retained, yet also how it needs to be extended. At the time Carroll and Sapon (1957) developed the MLAT, conceptions of memory did not include the separation between working and long-term memory. Yet in many ways, although the construct of phonemic coding was presented as a method of exploring how individual differences in sound processing impact on language learning, in fact, the construct brought together a number of abilities, not simply sound processing, but also something quite close to working memory as well as a capacity to analyse sound so that it can be retained. The implicit theory here is that:

- An immediate memory is important in verbal task performance
- The capacity to analyse sound in some way makes the sound easier to retain
- This analytic capacity may be connected to the ability to make connections between sounds and symbols
- Unfamiliar sound is particularly difficult to handle and perhaps most discriminating as the basis for aptitude testing

Three additional points need to be made about this analysis, and they concern the issue of analyzing unfamiliar sound. First, it has been assumed until now that this analytic ability is linked to the capacity to make connections between sound and symbol. One can retain the importance of analytic ability here but not necessarily accept that this analysis is only on the basis of sound-symbol connections. Second,

for sound-symbol connections to be available, it seems necessary to assume that the language material used in the test is sufficiently close to English to enable English-based sound-symbol associations to be relevant. This may not always be the case. Third, although the assumption is that it is *unfamiliar* sound that is important, little is done to explicate how unfamiliar can be defined, and what relevance it might have to aptitude test construction.

In that respect, proposals by Levelt and Wheeldon (1994) are relevant. They argue that during speaking, a mental syllabary is extremely important, and that speech production has an important syllable-based component. The syllabary contains information on high frequency syllables, and the access speeds to this syllabary, as well as the unitization of items within the syllabary facilitate speech production considerably. Levelt and Wheeldon (1994) also propose that the transitional probabilities within syllabaries are built up very early in life, and then are available to ease the task of speech production and processing throughout the lifespan. The learning is incidental, and not particularly apparent at the time when it is happening, but then becomes vital later for native-like processing.

The implication we can draw from this is that mental syllabaries in different languages will be different. Most generally, therefore, the second language learning task will be made more difficult if mental syllabaries for the L1 and the L2 diverge, and made easier if they resemble one another. In the first case, the sound structure of the target language may not come easily, while in the second, a great deal of transfer will be available. If this analysis of ease or difficulty of language learning is accepted, then it follows that if one is using non-word repetition tasks as part of a language aptitude battery, the way the nonwords are constructed will have a considerable impact on performance. If non words are constructed following a more

distant syllable or sound structure, the task will be more difficult. But more than that, if they are constructed on the basis of the syllabary and sound structure of the language to be learned, the difficulty of the test will also relate to the difficulty of the L2. Foreign language aptitude nonword repetition tests constructed in this way may therefore be selectively relevant and most predictively effective for specific languages. One could, of course, use this approach to unfamiliarity of sound structure simply to construct difficult tests. But more helpfully, it may be useful to construct tests which have maximum relevance, and allow more effective predictions for the learning of specific languages. It is also a method which should be relatively simple to implement.

This is a hypothesis for the future. We now have a relatively easy means for the construction of specific language oriented non-word repetition tasks as demonstrated by the particularly problematic case of Cantonese. What is required next is to conduct validation studies which pursue this insight.

Notes

- 1 The authors would like to thank Wen Zhisheng for comments on a previous draft of this article, and the useful input he has provided.
- 2 The authors would like to acknowledge the support of the Hong Kong Research Grants Council, through Grant No. 4707/05H which provide the funding to enable the research on which this article is based to be conducted.

References

- Baddeley A. (1986). *Working memory*. New York: Oxford University Press.
- Baddeley A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

- Baddeley, A., Gathercole S., & Papagno C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Carroll, J.B. (1965). The prediction of success in intensive foreign language training. In R.Glaser (Ed.), *Training, research, and education* (pp. 87-136). New York: Wiley.
- Carroll, J.B., & Sapon, S. (1959). *The Modern Languages Aptitude Test*. San Antonio, TX: Psychological Corporation.
- Clark, H., & Clark, E. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace Jovanovitch.
- Cronbach L. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.
- De Bot K., Lowie W., & Verspoor, M. (2006). *Second language acquisition: An advanced resource book*. London: Routledge.
- DeKeyser, R., & Juffs, A. (2005). Cognitive considerations in L2 learning. In E. Hinkel (Ed.), *Handbook of research second language teaching and learning* (pp. 437-454). Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2005). *Psychology of the individual language learner: Individual differences in second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum.
- Doughty, C. (2001). Cognitive underpinnings of a focus on form. In Robinson P. (Ed.), *Cognition and second language instruction* (pp. 206-257). New York: Cambridge.
- Gathercole, S. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition*, 23, 83-94.
- Gathercole, S. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513-543.
- Gathercole S., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, N.J.: Lawrence Erlbaum.
- Grigorenko, E.L., Sternberg R.J., & Ehrman, M. (2000). A theory based approach to the measurement of foreign language aptitude: the CANAL-F theory and test. *Modern Language Journal*, 84, 390-405.

- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25-38.
- Klein, W. (1986). *Second Language Acquisition*. Cambridge: Cambridge University Press.
- Lau, S. (1977). *A practical Cantonese-English dictionary*. Hong Kong: Government Printer.
- Levelt, W.J.M., & Wheeldon, L.R. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.
- Long M., & Robinson, P. (1998). Focus on form: theory, research, and practice. In C. Doughty and J. Williams J. (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Mitchell, R. & Miles, F. (2004). *Second language learning theories*. London: Hodder Arnold.
- Miyake, A., & Friedman, D. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A. F.Healy and L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp.339-364). Mahwah, N.J.: Erlbaum.
- Miyake. A., & Shah P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.
- Pawley A. and Syder F. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.), *Language and communication* (pp.191-225). London: Longman
- Petersen, C. & Al-Haik, A. (1976). The development of the Defense Language Aptitude Battery. *Educational and Psychological Measurement*, 36, 369-380.
- Pimsleur, P. (1966). *The Pimsleur Language Aptitude Battery*. New York: Harcourt, Brace, Jovanovic.
- Robinson P. (2002). Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstadt(1991). In P.Robinson (Ed.), *Individual differences and instructed language learning* (pp. 339-364). Amsterdam: John Benjamins.

- Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 631-678). Oxford: Blackwell.
- Sachs J.S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 2, 437-442.
- Schmidt, D. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 13, 206-226.
- Schmidt, R. (2001). Attention. In P. Robinson, (Ed.), *Cognition and second language acquisition* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, D. & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to Learning: Conversation in second language learning* (pp. 237-322). Rowley, Mass.: Newbury House.
- Schmitt, N. & Carter, R. (2004). Formulaic sequences in action: an introduction. In N. Schmitt, (Ed.), *Formulaic sequences* (pp. 1-22). Amsterdam: John Benjamins.
- Skehan, P. (1982). *Memory and motivation in language aptitude testing*. Unpublished Ph.D. thesis, University of London.
- Skehan, P. (1989). *Individual differences in second language learning*. London: Arnold.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning* (pp.69-94). Amsterdam: John Benjamins.
- Sparks, R.L. & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning difficulties: linking first and second language learning. *Modern Language Journal*, 77, 289-302.
- Stokes, S., Wong, A. M-Y., Fletcher P., & Leonard, L.B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research*, 49, 219-236.

- Swain, M. (1995). Three functions of output in second language learning. In Cook G. and B. Seidlehofer (Eds.), *Principle and Practice in Applied Linguistics* (pp. 125-144). Oxford: Oxford University Press.
- VanPatten, B. (1996). *Input Processing and Grammar Instruction*. New York: Ablex.
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52, 4, 755-803.

