Implicit Knowledge of Lexical Stress Rules: Evidence from the Combined Use of Subjective and Objective Awareness Measures

(forthcoming in *Applied Psycholinguistics*)

Ricky KW Chan

City University of Hong Kong and University of Hong Kong

Janny HC Leung

University of Hong Kong

Correspondence concerning this article should be addressed to Ricky KW Chan at rickychan0809@gmail.com.

Abstract

Despite the growing interest in the phenomenon of learning without intention, the incidental learning of phonological features, especially prosodic features, has received relatively little attention. This paper reports an experiment on incidental learning of lexical stress rules, and investigates whether the resultant knowledge can be unconscious, abstract and rule-based. Participants were incidentally exposed to a lexical stress system where stress location of a word is mainly determined by the final phoneme, syllable type and syllable weight. Learning was assessed by a pronunciation judgment task. Results indicate that participants were able to transfer their knowledge of stress patterns to novel words whose final phoneme was not previously encountered, suggesting that participants had acquired abstract and potentially rule-based knowledge. The combined use of subjective and objective measures of awareness in the present study provides a strong piece of evidence of the acquisition of implicit knowledge.

Keywords: Incidental learning, implicit knowledge, lexical stress, confidence ratings, source attribution, process dissociation procedure

Implicit Knowledge of Lexical Stress Rules: Evidence from the Combined Use of Subjective and Objective Awareness Measures

BACKGROUND

In the past few decades, a growing body of psycholinguistic research has focused on adults' capacity to learn without explicit instruction or metalinguistic information. This way of picking up information from the input without intention is generally referred to as incidental learning (Hulstijn, 2005). A typical incidental learning experiment often involves participants learning one stimulus aspect when paying attention to another (Hulstijn, 2008). For example, participants may learn some regularities of a grammar while performing a meaning-focused task. The term incidental learning should be distinguished from implicit learning, which refers to the above situation with an additional criterion that participants are unaware of the regularities to be learnt during the course of learning (Williams, 2009). The present study focuses on the nature of the linguistic knowledge acquired through incidental exposure in general.

Two major questions have been raised about incidental and implicit learning research. First, can incidental or implicit learning result in implicit knowledge? To address this, researchers need to assess participants' awareness of the acquired knowledge. However, most studies on incidental learning did not assess whether the knowledge acquired was implicit, and it is not always clear whether studies which claim to study implicit learning followed the necessary standards of design and procedure to assess participants' awareness of the resultant knowledge (Rebuschat, 2013). This is not surprising because, as noted by Leow (2015), the construct of awareness had not been methodologically and empirically addressed in the field of SLA until the latter part of the 90s. Sceptics maintain that the possibility of acquiring implicit knowledge has yet to be convincingly demonstrated, as no awareness measure developed so far is free from limitations. Given the multifaceted nature of the conscious state of knowledge, as we will show below, we argue that a combination of different types of awareness measures should be used, as far as possible, to assess implicit and explicit knowledge.

Second, what is the nature of the knowledge acquired through incidental or implicit learning? Reber (1989) argues that the resultant knowledge may be abstract and can be applied to perceptually different domains. In the psychology literature, several subsequent studies have provided evidence in support of this view (e.g. Altmann et al., 1995; Goschke & Bolte, 2007; Rohrmeier et al., 2012; Scott & Dienes, 2010; Tunney & Altmann, 2001). On the other hand, some researchers demonstrated that implicit knowledge consists of merely memorized chunks or details of particular exemplars (e.g. Brooks & Vokey, 1991; Dulany et al., 1984; Jamieson & Mewhort, 2011; Perruchet & Pacteau, 1990; Pothos, 2007; Pothos & Bailey, 2000). In the realm of second language acquisition, while transfer of abstract knowledge to new items has been shown in the learning of syntax (e.g. Rebuschat, 2008; Williams and Kuribara, 2008), the acquisition of implicit abstract knowledge of L2 phonology, especially in the suprasegmental domain, has received relatively little attention (c.f. Chan and Leung, 2014 for an exception). Research in this area will contribute to the development of a comprehensive model of implicit and explicit L2 knowledge.

The goal of this paper is to contribute to our understanding of language learning in two ways: 1) we aim to explore the possibility of acquiring abstract and potentially rule-based knowledge of stress patterns incidentally; and 2) we provide a strong piece of evidence of implicit knowledge using multiple measures of awareness. Let us first provide an overview of whether incidental or implicit learning may result in abstract rule knowledge, followed by a discussion on how conscious knowledge and unconscious knowledge may be measured. Previous work on the acquisition of implicit knowledge of L2 word stress patterns will also be reviewed.

Can incidental or implicit learning result in abstract rule knowledge?

A theoretical question in incidental or implicit learning research concerns the nature of the resultant knowledge. In the artificial grammar learning (AGL) paradigm where implicit learning was originally studied by Reber (1967), participants were presented with sequences of letters (e.g. VXVS) generated by a finite state grammar and were told to memorize them as part of an experiment on rote memory. Implicit learning was demonstrated when participants achieved above-chance performance in the subsequent grammaticality judgment task but were unable to report the underlying rules used to generate the letter strings. Reber (1989) stated that participants in the AGL experiment had acquired abstract, and potentially rule-based, knowledge of the underlying structure. In another widely used implicit learning paradigm called serial reaction time (SRT) tasks (Nissen & Bullemer, 1987), a stimulus (e.g. a dot) appears on one of four or six locations on the screen, and participants respond by pressing a corresponding button as quickly as possible. While the task appears to be testing participants' reaction time, unbeknownst to participants there are rules governing the sequence of the location of the stimulus. Learning of the underlying rules is usually demonstrated when participants respond faster when the sequence follows the rules than when the sequence violates the rules.

However, it has been argued that implicit knowledge reported in AGL and SRT paradigms is based on knowledge of chunks (e.g. Dulany, Carlson, & Dewey, 1984; Johnstone & Shanks, 2001; Kinder & Assmann, 2000; Perruchet & Pacteau, 1990; Servan-Schreiber & Anderson, 1990) or details of particular exemplars (e.g. Brooks &

Vokey, 1991; Jamieson & Mewhort, 2011; Jamieson & Hauri, 2012; Pothos, 2007) instead of abstract rule induction. For example, the test item VXVS contains the bigrams VX, XV and VS and the trigrams VXV and XVS. Dulany et al. (1984) replicated Reber's study and found that participants' above-chance performance on a grammaticality judgment task in an AGL experiment can be attributed to memorized fragments of the letter stings. Perruchet and Pacteau (1990) also found similar test performance between participants who had only been trained on grammatical pairs of letters and those who had exposure to the complete strings. In a similar vein, Perruchet (1994) found that, in a sequence learning task, participants were only sensitive to the similarity between old and new sequences and there was no evidence of learning of the underlying rule. These results suggest that grammaticality decisions in these experiments may not be based on abstract rule knowledge as Reber originally claimed. Still, in the case of language learning, we often assume that the grammatical representations learners internalize can be applied to new stimuli which have no surface similarity to previous utterances (e.g. we know the sentence *Colourless green ideas sleep furiously* is syntactically well-formed) (Williams, 2009). In AGL research, the implicit abstraction issue has been examined by testing whether participants can transfer their knowledge to different sets of letters (e.g. VXVS in training may correspond to ABAC in test) or different modalities (e.g. letter strings in training and tone sequences at test), where the underlying grammar is the same. Altmann et al. (1995), for example, found that participants' performances on classifying strings in a different modality was above chance level, although they performed less well than in the same modality condition. This suggests that at least part of the knowledge acquired was abstract and potentially rule-based, rather than merely consisting of memorized fragments or chunks. Several other studies also support this view (Goschke & Bolte, 2007; Jiang et al., 2012; Knowlton & Squire, 1994, 1996; Rohrmeier, Fu, & Dienes, 2012; Rohrmeier, Rebuschat, & Cross, 2011; Scott & Dienes, 2010; Tunney & Altmann, 2001; Turk-Browne & Scholl, 2009). Evidence of transfer of abstract knowledge to novel stimuli has been shown in the incidental learning of syntax (e.g. Rebuschat, 2008; Williams and Kuribara, 2008). Still, relatively few studies have focused on the prosodic domain. The present study thus aims to examine whether people can acquire and transfer abstract and potentially rule-based knowledge of lexical stress assignment.

Measuring implicit and explicit knowledge

A key methodological challenge in the study of implicit and explicit knowledge lies in how awareness should be operationalized and assessed. A distinction is often drawn between "subjective" and "objective" awareness measures: subjective measures generally require participants to report what they think they know, whereas objective measures assess participants' knowledge based on their performance or behaviour. Here we review three sets of awareness measures commonly used in cognitive psychology that have recently been applied to SLA research: retrospective verbal reports, confidence ratings and source attribution, and process dissociation procedure.

Retrospective verbal reports. A commonly used subjective measure is retrospective verbal reports, which involve prompting participants to verbalise any patterns that they have noticed after the learning and testing tasks. Verbal reports have often been used in SLA research (e.g. Williams, 2005; Leung & Williams, 2011, 2012, 2014). Knowledge is considered implicit if participants show a learning effect (e.g. above-chance performance in a judgment task) but fail to report any knowledge of the learning target. However, the validity of verbal reports has been challenged on several grounds: its insensitivity to low-confidence knowledge, the dissociation between the acquired knowledge and its verbalizability, potential memory decay, and the fact that the knowledge reported may not have contributed to the performance on the measure of learning (see Rebuschat, 2013; Shanks & St. John, 1994 for reviews). Still, although retrospective verbal reports as an awareness measure may seem to be insensitive and incomplete, they may be sensitive to participants' verbalizable knowledge and provide some insights into what participants have learnt.

Confidence Ratings and Source Attribution. According to the higher order thought theory (Rosenthal, 2005), an experience is conscious when there is a higher order thought asserting that we have that experience. For example, a conscious experience of green is composed of a representation of green in the visual system, along with a higher order thought of the experience of green (a meta-representation). Based on the higher order theory of consciousness, Dienes and his collaborators (Dienes, 2004, 2008; Dienes et al., 1995; Dienes & Berry, 1997; Dienes & Scott, 2005) have proposed the use of trial-by-trial confidence ratings and source attributions to assess participants' conscious state. Confidence ratings involve asking participants to report how confident they were when making their decision. For example, with a simple confidence scale consisting of two options "guess" or "know", "guess" indicates that the person's judgment had no firm basis, whereas "know" indicates that the judgment was based on some knowledge. Dienes et al. (1995) give two criteria for which confidence ratings data can assess conscious knowledge. First, if accuracy of participants' decision is above chance when participants are believed to be guessing, they can then be said to be using implicit knowledge. Dienes et al. called this the *guessing criterion*. Second, knowledge is unconscious when there is no relationship between participants' accuracy and their confidence. This criterion, introduced by Chan

(1992), was labelled *zero correlation criterion* by Dienes et al. (1995). However, the use of confidence ratings has been criticized for assessing only judgment knowledge (knowledge about whether a particular test item should be classified as the same or different from the training items), but not structural knowledge (knowledge of the structure of the training items such as fragments or rules), and the two kinds of knowledge can in theory be separated. In light of the criticism on the confidence ratings task, Dienes and Scott (2005) subsequently developed source attribution as a way to assess participants' structural knowledge. In their study, participants were asked to report the basis of their judgment—pure guess, intuition, memory of part or all of the training items, or a rule/rules that they can state. They argue that the guess or intuition attribution indicates unconscious structural knowledge, whilst the memory or rule attribution conscious structural knowledge. In short, these two subjective measures assess the existence of relevant higher order thought of participants' judgement knowledge and structural knowledge.

However, as noted by Overgaard et. al (2010), the field of cognitive science has generally disregarded subjective data and preferred objective data as the major source of evidence for participants' conscious state. One potential problem of using confidence ratings and source attribution as an awareness measure is that participants may set their own criterion for reporting knowledge. For example, more conservative participants may state that they are guessing on their grammaticality judgments unless they are absolutely sure, while more liberal participants may consistently report high levels of confidence even at the slightest intuition. Also, there is no guarantee that participants would have reported all relevant judgment and structural knowledge, and thus these awareness measures may not be sensitive to all relevant knowledge participants have, failing to fulfil the sensitivity criterion for the test of awareness (Shanks and St. John, 1994). Still, confidence ratings and source attributions have been shown to be more sensitive in detecting low-confidence conscious knowledge than verbal reports (e.g. Ziori and Dienes, 2006), and have recently been used in SLA research (e.g. Grey, Williams & Rebuschat, 2014, 2015; Hamrick & Rebuschat, 2012; Rebuschat et al., 2013; Rebuschat & Williams, 2012; Rogers, & Rebuschat, 2016).

Process dissociation procedure (PDP). As an objective measure of awareness, PDP was first proposed by Jacoby (1991) to disentangle the contribution of implicit and explicit knowledge based on participants' behaviour. The basic principle is to design two tasks: one in which implicit and explicit knowledge act in concert, and one in which implicit knowledge interferes the contribution of explicit knowledge to performance (Jacoby et al., 1993). The amount of explicit knowledge acquired can be estimated by the difference in the performance of the two tasks. PDP avoids the problem of process purity and takes into account the fact that both implicit and explicit knowledge contribute to any task performance (Dunn & Kirsner, 1989; Jacoby, 1991). Destrebecqz and Cleeremans (2001) adapted PDP in the SRT task: after completing the SRT task, participants were informed that the presentation of the visual stimuli followed a repeating pattern, and were instructed to complete free-generation tasks under two conditions: 1) generate as much of the training sequence as they can (inclusion condition); and 2) generate a different sequence (exclusion condition). According to the Global Workspace theory (Baars, 2003), when knowledge becomes conscious, the possibility for voluntary control of performance is opened up. Participants who possessed some explicit knowledge tended to follow the sequence in the inclusion condition but not in the exclusion condition; a difference between inclusion and the exclusion performance indicates top-down processing and thus explicit knowledge. By contrast, participants with no explicit knowledge tend to perform equally well in both inclusion and exclusion tasks (Curran, 2001).

Although PDP has widely been used in the field of cognitive psychology, few attempts have been made to apply PDP in SLA research (c.f. Chan & Leung, 2014 for an exception). A major advantage of using PDP is the possibility of assessing awareness based on objective data. Still, the use of PDP hinges on the assumption that conscious knowledge is reflected in performance, which may not always be warranted. Some believe that consciousness as a subjective experience cannot be observed from the outside and conscious knowledge may be separate from performance (Dienes, 2008; Overgaard et al., 2010).

It should be clear from the discussion above that none of the above awareness measures is free from limitation. Nonetheless, these awareness measures are sensitive to different aspects of consciousness—verbal reports capture verbalizable knowledge; confidence ratings and source attribution are sensitive to low-level subjective knowledge represented by corresponding higher order thought; and PDP allows us to objectively assess the degree to which knowledge is subject to controlled processes. While a single awareness measure may not be able to capture the seemingly multi-faceted nature of consciousness, a possible solution is to employ more than one awareness measures which complement one another to deal with the complex nature of the phenomenon of (un)consciousness (Rebuschat, 2013; Seth, 2008). Accordingly, in the present study on the acquisition of L2 lexical stress patterns, we combined subjective and objective awareness measures in a bid to provide a stronger piece of evidence for the possibility of acquiring implicit knowledge.

Acquiring Implicit Knowledge of Lexical Stress Patterns

Lexical stress plays an important role in organizing the speech stream. Knowledge of lexical stress patterns plays a key role in various language processing tasks such as parsing the speech stream (e.g. Trubetzkoy, 1969) and memorising novel words (Bell, 1977; Cutler, 1986). However, only a few studies have focused on whether implicit knowledge of stress rules can be acquired. Bailey et al. (1999) conducted a cross-linguistic study on the learnability of rhythm patterns. They found that knowledge of complex stress patterns can be generalised even after only brief exposure, and that typologically less common stress patterns were easier to learn. However, they substituted pitch patterns for stress patterns; it remains questionable whether the findings can be generalized to the learning of linguistic stress patterns as f0 is not the only acoustic correlate of linguistic stress. Moreover, they claimed that participants' knowledge was implicit because they reported no awareness of the learning targets retrospectively, but this assumption may not be warranted as discussed above.

Zellers, Post and Williams (2011) investigated the incidental learning of simplified Spanish-based stress patterns: /s/-final or words ending in an open syllable stress on the penultimate syllable, and consonant-ending words (other than /s/) stress on the final syllable. Native English speakers were trained with a short-term memory task in which they heard sequences of words and then repeated aloud. In the testing phase, they were then presented with novel words and asked if they had heard the words during training. They postulated that if participants had learnt something about the stress patterns, they would be more likely to state they had heard a novel word when the word follows the target stress patterns. However, the results might have been confounded by the fact that, when stating they had or had not heard a particular word during training, participants' judgments might not have been based on their knowledge of the target stress patterns, but on other aspects of the word (e.g. its phonemes). Also, there was no clear evidence showing that participants' knowledge was abstract and rule-based. Furthermore, they assessed whether participants' knowledge of the target patterns was implicit by retrospective verbal reports only. Although participants were unable to verbalize the rule they had learnt, participants with low level of awareness or confidence might have left undetected.

Chan and Leung (2014) studied implicit learning of simplified L2 stress patterns which concern the mapping between the final phoneme(s) in a word and stress assignment: *o*-final words have penultimate stress and *ar*-final have word-final stress. The study demonstrated that participants acquired implicit knowledge of one-to-one phoneme-to-stress mapping, but the possibility of acquiring abstract unconscious knowledge of stress patterns (e.g. connections between syllable weight and stress location) has yet to be explored. Besides, while participants' awareness was assessed by retrospective verbal reports and the PDP, it is still possible that participants with low-confidence knowledge of the target stress patterns went unnoticed.

RESEARCH QUESTIONS

Given that the possibility of acquiring unconscious abstract L2 knowledge remains controversial, the present study has two specific research questions:

1) Could incidental learning result in abstract and potentially rule-based knowledge of lexical stress patterns?

In the experiment described below, the learning targets involved the mapping between stress assignment and the more abstract phonological categories of vowels and consonants. If incidental learning of stress rules resulted in abstract representation, independent of the items in the training set, participants should be able to transfer their knowledge stress assignment for novel words.

2) Could incidental learning lead to unconscious knowledge of lexical stress patterns?

Most previous studies on implicit/incidental learning used only one (type) of awareness measure which, as we argue above, may be insufficient to capture the multidimensional nature of the phenomenon of (un)consciousness. In the present study, we aimed to provide a more comprehensive assessment of implicit knowledge with the combined use of verbal reports, confidence ratings, sources attribution, and PDP. If participants exhibit no awareness of the target rules from all of these measures but nevertheless show learning effects, we are more confident to conclude that incidental learning may lead to unconscious knowledge of lexical stress patterns than if we had relied on a single measure.

METHOD

Learning Targets

The learning targets were stress[i] rules which are determined by the final phoneme, as illustrated below:

1. Consonant-final words stress on the final syllable (e.g., *fe**lol*** and *ce**rroz***).
2. Vowel-final words stress on the first syllable (e,g, ***pa**to* and ***bo**na*).

Only disyllabic words were used in the entire experiment. As such, the consonant-final words and vowel-final words also differ in terms of syllable weight (light vs. heavy) and syllable types (closed vs. open) in the last syllables. The resultant stress system resembles a trochaic, weight-sensitive system with right-edge footing. As a matter of fact, such a stress system is very common in natural languages (see e.g. Hayes 1995).

Stimuli

All the stimuli were generated by the MBROLI speech synthesizer using a diphone database of a European Spanish speaker (es1)[1] (Dutoit et al., 1996), and thus the phonetic realisations of the stimuli were based on European Spanish. Spanish was chosen in a bid to 1) minimize effects of prior linguistic knowledge and proficiency for the L1 Cantonese L2 English participants; 2) ensure that no lexical meaning was previously associated with the novel word forms; and 3) control for vowel quality in stressed and unstressed syllables, as stress contrast does not lead to vowel quality contrast in Spanish.

All the disyllabic real words and nonce words consist of four or five phonemes (vowel-ending words and consonant-ending words respectively; except for *h*-initial words). All the words used in the training, which complied with the target stress rules, were concatenations of phonemes listed below:

1. First phoneme: /b/, /d/, /f/, /g/, /k/, /ʎ/, /m/, /n/, /p/, /s/, /t/, /∅/.

2: Second phoneme: /a/, /e/ or /o/.

3. Third phoneme: /b/, /d/, /g/, /k/, /l/, /m/, /n/, /p/, /ɾ/, /r/, /s/, /t/.

4. Fourth phoneme: /a/, /e/ or /o/

5. Fifth phoneme (consonant-ending words only): /ɾ/, /l/, /θ/

(e.g. fane /'fane/, llaner /'ʎa'neɾ/)

Two additional phonemes, /d/ and /x/, were used in the testing phase (see "Testing Phase" in the Procedure section below). Some of the consonants (/ʎ/, /∅/, /r/, /x/) do not exist in the phoneme inventory of either Cantonese or English. The same set of vowels was used for both stressed and unstressed syllables to control for vowel quality.

[1]https://tcts.fpms.ac.be/synthesis/mbrola/mbrcopybin.html

The manipulations of the stimuli were as follows: each word starts with a 50ms silence and ends with another 50ms silence. The f0 peak and duration of the vowels were set at 100Hz and 90ms for unstressed vowels and 116Hz and 120ms for stressed vowels (Face, 2005), and the duration of the consonants can be found in Appendix I. The specific manipulations of f0 are described as below:

Words with initial stress. A flat f0 of 116Hz was placed on the first syllable until the beginning of the second syllable. f0 then lowers to the f0 target of 100Hz at 25% of the second vowel, and further lowers to the f0 target of 90Hz at the end of the second vowel. The speech synthesizer automatically dropped f0 information when synthesising voiceless segments. This is illustaed in Figure 1.

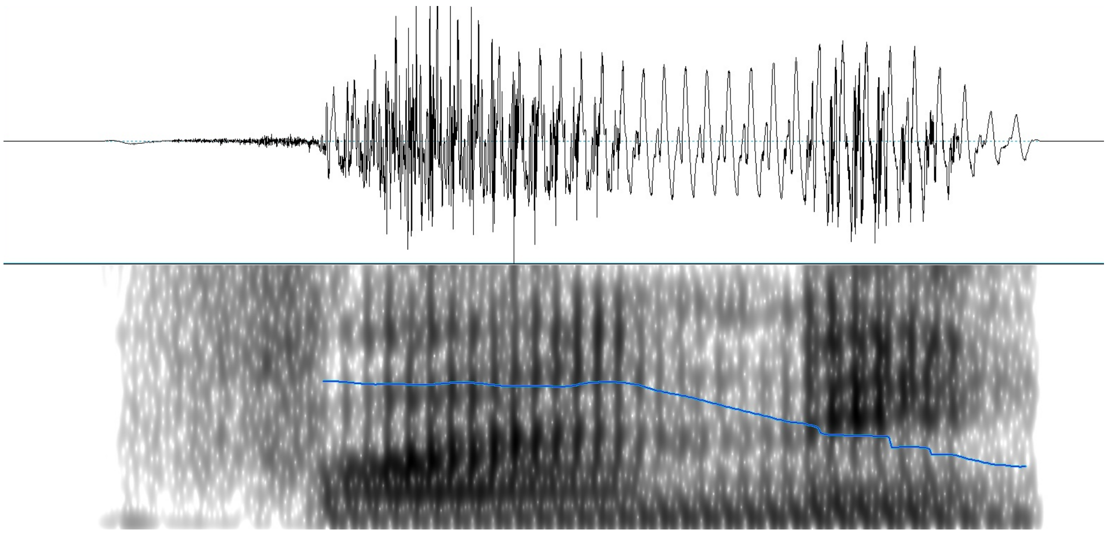

Figure 1. F0 manipulation of the words with initial stress (e.g. “fane” /ˈfane/).

Words with final stress. A flat f0 of 100Hz was placed on the first syllable until the beginning of the second syllable. f0 is then raised to the f0 target of 116Hz at the centre of the second vowel, and then lowers to the f0 target of 90Hz at the end of the second vowel. This is illustrated in Figure 2.

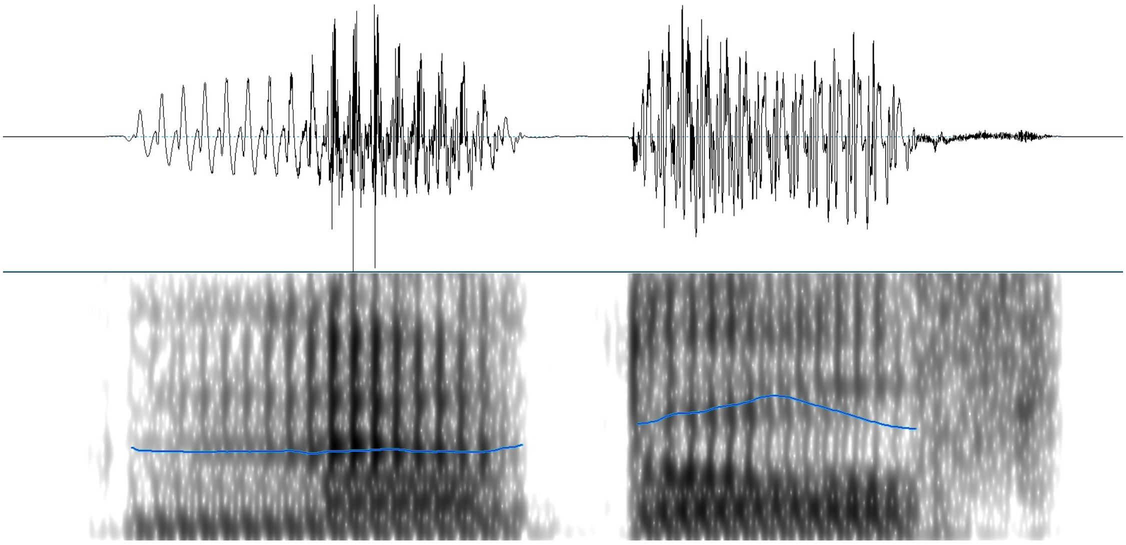

Figure 2. F0 manipulation of the words with final stress (e.g. “natoz” /naˈtoθ/).

The frequency of each phoneme in all positions was counter-balanced for all vowel-ending words and consonant-ending words used in the experiment. This served to prevent participants from relating stress assignment other than the nature of the last phoneme.

Subjects

Ninety L1 Cantonese L2 English (intermediate to upper intermediate, based on self-report) undergraduates participated in the study. Sixty-five (31 males and 34 females, M_{age} = 21.1) participants were assigned to the experimental group and twenty-five (12 males and 13 females, M_{age} = 21.7) to the control group. Sixty-three of them reported having taken at least one university course related to linguistics. None of them reported any knowledge of Spanish or other languages that have lexical stress.

Participants’ existing linguistic systems may facilitate learning of the above novel lexical stress patterns. Their L1, Hong Kong Cantonese, is a tone language which uses distinctive pitch patterns to distinguish word meaning. Cantonese contrasts three level tones (high [55], mid [33] and low [22]), two rising tones (high [25] and low [25]) and a low falling tone [21][2] (Bauer, 1997). The primary perceptual correlate of Cantonese tone is pitch (Bauer, 1997), which is also one of the major perceptual correlates of lexical stress (Morton & Jassem, 1965). The rich pitch contrasts in their tone system may aid their perception of stress contrast. When learning L2 lexical stress, Cantonese speakers tend to make use of tone contrast to represent lexical stress contrast. For instance, Cantonese speakers consistently assign high level tone to stressed syllables and lower level tones to unstressed syllables in English (Chao, 1980; Luke, 2000). Data from Cantonese loanwords also exhibit similar patterns: Cantonese speakers assign high level one to stressed syllables, mid-level tone to unstressed syllables and mid-low level tone to epenthetic syllables in English donor words (Lai et al., 2011). Studies on lexical stress perception also corroborate these findings. In Chan's (2007) study of word stress perception, he manipulated the fundamental frequency (f0), duration, and spectral balance of <bebe> logatomes; L1 Cantonese L2 English participants were instructed to judge the position of lexical stress of the logatomes which were embedded in a carrier sentence. He found that f0 alone accounted for most of the listeners' responses (79%) in stress perception; by contrast, duration and spectral balance only played a minimal role. Similarly in Tong et al. (2015), L1 Cantonese L2 English participants were asked to discriminate lexical stress patterns in psuedoword CVCV sequences in the AXB discrimination paradigm. They found that average F0, F0 onset, F0 offset, F0 general slope, duration, intensity, and spectral balance are important acoustic cues to distinguish among stress patterns.

Therefore, we expect our participants to use the cues we manipulated for stress perception.

---

[2] The number in [] represents relative pitch height with reference to a speaker's pitch range, ranging from 1 (lowest) to 5 (highest) (Chao, 1947).

On the other hand, participants' L2, English, has a lexical stress system which is quantity sensitive: heavy syllables (when the rhyme is a tense vowel, a diphthong or closed by a consonant) tend to attract lexical stress (Hayes, 1995). At the same time, in English feet are trochaic (left-head) and are iteratively built from right to left with extrametrical final syllables, accounting for the preference for non-final stress in English (Hammond, 1999; Hayes, 1995). Whilst participants may prefer initial stress for disyllabic words, they are also expected to be sensitive to the relationship between syllable weight and stress location. Therefore, we expect them to show no preference for either vowel- or consonant-final words.

Procedure

All instructions were presented in both Chinese and English and the experiment was administrated on a computer using E-prime. Participants in the experimental group went through all the tasks described below whereas those in the control group only completed the pronunciation judgment task.

Training phase. To disguise the purpose of the experiment, participants were told that the experiment aimed to study how people learn words. In each trial, participants were visually and aurally presented with a word (Figure 3). Participants were instructed to listen to the word and repeat it aloud. No definition or translation of the word was provided. This encouraged them to pay attention to the pronunciation of the letters and stress location. According to Schmidt (1990, 2001, 2010), "noticing" (conscious registration of attended input) is necessary in SLA, but not "understanding" (a higher level of awareness that involves generalisations across instances, such as knowledge of rules and metalinguistic awareness). Providing the input both visually and aurally and asking them to repeat promoted "noticing" of the pronunciation of the phonemes and stress assignment of the word. However, no explicit information was provided about the mappings between the ending phoneme and stress placement, which were the learning target of the experiment.

Madel

Figure 3. The visual presentation of a sample trial in the training phase.

A set of 36 words (Table 1) were presented in a random order; the whole set was repeated four times to form 144 trials.

Vowel-ending			**Consonant-ending**		
a-final	**o-final**	**e-final**	**r-final**	**l-final**	**z-final**
Beba /'beba/	Gobo /'gobo/	Coge /'koge/	Bator /ba'toɾ/	Bogel /bo'gel/	Cerroz /se'roθ/
Bona /'bona/	Navo /'nabo/	Dade /'dade/	Coder /ko'deɾ/	Debal /de'bal/	Gapez /ga'peθ/
Cepa /'sepa/	Pato /'pato/	Fane /'fane/	Llaner /'ʎa'neɾ/	Domal /do'mal/	Hocaz /o'kaθ/
Doca /'doka/	Seco /'seko/	Mete /'mete/	Penar /pe'naɾ/	Felol /fe'lol/	Natoz /na'toθ/
Hara /'aɾa/	Sorro /'soro/	Tome /'tome/	Socor /so'koɾ/	Madel /ma'del/	Tobaz /to'baθ/
Llada /'ʎada/	Telo /'telo/	Vese /'bese/	Tevar /te'baɾ/	Sasol /sa'sol/	Verez /be'reθ/

Table 1. Items used in the training phase (transcribed phonemically)

Testing phase. Participants were tested on the stress rules by a two-alternative-forced-choice pronunciation judgment task. In each trial, participants pressed relevant keys to listen to two words (shown as “word 1” and “word 2”; see Figure 4) and chose the one that “sounded better” to them; this, when compared with accuracy judgment such as “choose the correct one”, encouraged the use of intuition and discouraged rule search.

Figure 4. The visual presentation of a sample trial in the testing phase.

18 novel nonce words, half vowel-final (*a*, *e*, *o*) and half consonant-final (*r*, *l*, *z*), served as critical items (Table 2). A further 12 nonce words which end in a novel vowel (/i/ or /u/) or consonant (/d/ or /x/) were included as extension items (Table 3). Sound pairs for the critical and extension items differed only in stress assignment. In this way, only by knowing the target stress rules could they choose the correct answers. If participants possessed abstract knowledge of the target stress patterns, they should be able to apply their knowledge to novel stimuli whose last phoneme was either encountered in the training phase (critical items) or unseen and unheard (extension items). 18 of the training items, half vowel-final and half consonant-final, were also included as fillers so that participants were less likely to be aware of the purpose of the test.

Vowel-ending			**Consonant-ending**		
a-final	**o-final**	**e-final**	**r-final**	**l-final**	**z-final**
Dada /'dada/	Goto /'goto/	Cebe /'sebe/	Decar /de'kaɾ/	Mebel /me'bel/	Cepez /se'peθ/
Moda /'moda/	Llemo /'ʎemo/	Sage /'sage/	Llager /'ʎageɾ/	Savol /sa'bol/	Gadoz /ga'doθ/
Teca/'teka/	Savo /'sabo/	Tope /'tope/	Tomor /to'moɾ/	Sotal /so'tal/	Todaz /to'daθ/

Table 2. Critical items used in the testing phase

Vowel-ending		**Consonant-ending**	
i-final	**u-final**	**d-final**	**j-final**
Llepi /'ʎepi /	Dotu /'dotu/	Seded /se'ded/	Gotej /go'tex/
Gomi /'gomi/	Sacu /'saku/	Camod /ka'mod/	Llecaj /ʎe'kax/
Cabi /'kabi/	Tedu /'tedu/	Tobad /to'bad/	Dapoj /da'pox/

Table 3. Extension items used in the testing phase

In each trial of the testing phase, we also assessed the conscious status of participants' judgment knowledge and structural knowledge with confidence ratings and source attribution. After each pronunciation judgment, they were first instructed to indicate how confident their judgment was on a binary scale, explained to the participants as follows: "guess"—you are making a completely random guess; and "know"—you have some confidence in your choice. Binary confidence ratings were adopted as they are more sensitive to low levels of awareness than continuous confidence ratings (Tunney & Shanks, 2003; Tunney, 2005).

Participants then stated the basis for their decision as either "guess", "intuition", "memory" or "rules", which were defined to them as follows: "guess"—you have no idea and are making a random guess; "intuition"—you think your choice is right but have no idea why; "memory"—your choice was based on a recollection of any training item; and "rules"—your choice was based on one or more rules or partial rules that you can state. Dienes and Scott (2005) argue that the guess and intuition attributions reflect unconscious structural knowledge, whereas memory and rule attributions reflect conscious structural knowledge.

Inclusion-exclusion tasks. Participants were informed at the beginning of this part that there were rules governing the stress assignment of the words presented. To illustrate the concept of "stress assignment", participants were given a minimal pair in English "IMport" and "imPORT" which differ in their stress location. The inclusion-exclusion tasks, adapted from Jacoby (1991) and Destrebecqz and Cleeremans (2001), required participants to pronounce 36 new nonce words (Table 4) under two conditions: 1) for the first half of the words, pronounce them in a way that follows the underlying stress rules (inclusion condition); and 2) for the second half of the words, pronounce them in a way that does not follow the rules (exclusion condition). The inclusion task encouraged implicit and explicit knowledge to act in concert, whereas the exclusion task in opposition. A dot was given to indicate syllabification (Figure 5), making it clear to participants that all words consisted of only two syllables. Words were presented in a random order, and the frequencies of words with each phoneme ending were the same in the two conditions. Their voices were recorded.

Go.to

Figure 5. A sample trial in the inclusion-exclusion tasks.

Vowel-ending			**Consonant-ending**		
a-final	**o-final**	**e-final**	**r-final**	**l-final**	**z-final**
Ho.na	So.to	Ta.re	To.bar	Vo.sal	De.rraz
Ce.ba	Ga.lo	Lle.de	Ca.mer	Ba.pel	Fo.gez
No.ca	Ba.vo	Me.te	Se.nor	Pe.col	Da.doz
Ga.ba	Me.no	Va.de	Bo.ver	Ne.bol	Da.coz
Te.la	Co.po	Se.ge	Ce.rror	Fo.tal	Te.maz
Do.sa	Be.to	So.ne	Lle.car	Pa.rrel	Ho.dez

Table 4. Items used in the inclusion-exclusion task

Retrospective verbal reports. Participants were first asked to report any patterns they had noticed about the words they have learnt. As they were already informed that there are stress rules for the words in the inclusion-exclusion tasks, they were then encouraged to make as many guesses as possible about the underlying stress rules.

RESULTS

Classifying Aware and Unaware Participants

Participants in the experimental group were first classified into aware and unaware groups based on verbal reports, confidence ratings and inclusion-exclusion tasks.

Retrospective verbal reports. Most of the participants had no idea that there were rules governing the location of stress and were surprised when told so. Participants who made no guess at all or made guesses which did not overlap with the target stress rules (e.g. “usually stress the first syllable”; “stress is related to part of speech or meaning”) were classified as “unaware”. On the other hand, 4 participants were able to verbalize the whole target stress rules. 7 other participants reported knowledge that overlapped with the target rules: 3 guessed stress was related to the length of the word; 3 noticed that “z”, “l” and “d” were “heavily pronounced” and one stated the reverse stress rules. All of these participants were classified as “aware” as they were able to report knowledge that at least partially overlapped with the target stress rules. A summary is provided in Table 5. The use of such a strict criterion in our classification ensured that participants with any detectable level of awareness would not enter the unaware group.

Classified as "Aware"		
Verbal Reports	***N***	**Example**
Able to verbalise the whole target rules	4	- Initial stress for V-final words; final strss for C-final words
Reported knowledge of the target rules	7	- "z", "l", "d" are "heavily pronounced" - Stress is related to the length of the word
Classified as "Unaware"		
Made no guess or made guess with no overlapping with the target rules	54	- Stress location is related to part of speech/meaning/gender case - Stress is usually on the 1st syllable - Stress is usually on the 2nd syllable - Some sounds tend to attract stress - Initial "r", "n" or "c" may attract stress

Table 5: Summary table of the classification of "aware" and "unaware" participants based on retrospective verbal reports.

Confidence ratings. According to the zero-correlation criterion, knowledge is unconscious when accuracy does not correlate with confidence. The Chan difference score (Chan, 1992) of each participant was computed to determine whether they possessed conscious judgment knowledge. For a binary confidence rating, the score is calculated as the proportion of "know" responses which were correct (*hit*) minus those which were incorrect (*false alarm*). Participants with a positive score were classified as "aware", as they possessed conscious judgment knowledge as measured by the zero correlation criterion; those scored 0 or below were classified as "unaware".

Inclusion-exclusion tasks. We only analysed participants' assignment of stress in the novel words. We did not analyse other aspects of their pronunciation (e.g. how they realised the segments) since those were not the focus on the present study. Participants' assignment of stress was mainly determined auditorily; whenever uncertainties arose, stress assignment was determined acoustically in *Praat* based on the relative f0 and duration of the two syllables. The syllable with a higher f0 and longer duration was labelled as a stressed syllable. A trial was considered correct when stress was placed on the correct syllable in inclusion condition and the incorrect one in exclusion condition, regardless of how the segments were pronounced. The contribution of explicit knowledge was determined by subtracting the number of correct responses in the exclusion task from that in the inclusion task. Equal performance in the two tasks would indicate unconscious knowledge, whilst a positive or a negative score suggests sensitivity towards the underlying patterns.

Based on the three awareness measures, 40 out of 65 participants were classified as "aware"; the specific breakdown is presented in Table 6 (note that some participants displayed awareness in more than 1 awareness measure). The other 25 participants were classified as "unaware".

Aware Participants as Identified by	***N***
Verbal reports only	1
Confidence ratings only	7
Inclusion-exclusion tasks only	10
Both verbal reports and confidence ratings	0
Both verbal reports and inclusion-exclusion tasks	4
Both confidence ratings and inclusion-exclusion tasks	12
All three measures	6
Total number of aware participants	**40**

Table 6. Number of aware participants whose awareness was captured by different (combinations of) awareness measures. Note that some participants revealed their awareness of the target patterns in more than one measure.

Several observations can be made here. First, while confidence ratings and inclusion-exclusion tasks captured roughly similar numbers of aware participants (25 and 32 respectively), verbal reports identified relatively fewer aware participants (11). Also, with only one exception, participants whose conscious knowledge was captured by verbal reports were also identified as aware by confidence ratings or inclusion-exclusion tasks. On the contrary, most participants who were classified as aware based on confidence ratings or inclusion-exclusion tasks (or both) were not identified by verbal reports (29 out of 44). This suggests that verbal reports were only able to identify participants with a relatively high level of awareness, and thus constitute a less sensitive measure of awareness.

Interestingly, of the six participants classified as "aware" by all three measures, five were able to verbally report the whole target rules, and the remaining one mentioned that stress location is related to the length of word. This suggests that when participants became highly or fully aware of the target rules, their awareness would likely be reflected in all three awareness measures.

Performance in the pronunciation judgment task

Table 7 shows participants' performance on the critical items and the extension items. Figures 6 and 7 show the individual performance for the critical items and extension items respectively. All the analyses below were based on one-sample two-tailed t test unless otherwise specified. The overall accuracy of the 65 participants in the experimental group on the critical items was significantly above chance (50%), $t(64) = 8.26$, $p < 0.001$, $d = 1.46$, suggesting that L2 stress patterns can be acquired as a result of incidental exposure. Both the aware and the unaware groups achieved above-chance accuracy, $t(39) = 6.84$, $p < 0.001$, $d = 1.55$ and $t(24) = 4.88$, $p < 0.001$, $d = 1.41$ respectively. However, it should be noted that the "aware" participants did not constitute a homogenous group but may vary in their degree of awareness. It would be simplistic to compare the performances of the "aware" group and the "unaware" group and assume that they represent the contrast between the contribution of explicit and implicit knowledge. On the other hand, the control group, which completed the pronunciation judgment task without training, did not achieve above-chance accuracy on critical items, $t(24) = 1.79$, $p = 0.08$. Importantly, independent two-tailed t test shows that the unaware participants performed significantly better than the control group, $t(48) = 2.92$, $p = 0.0053$, $d = 0.83$, indicating that the implicit knowledge of the unaware participants resulted from the training.

Critical Items	**N**	**Mean (%)**	**SD (%)**	**SE (%)**
Overall	65	61.6	10.85	1.35
Aware	40	62.8	11.82	1.87
Unaware	25	58.4	8.65	1.73
Control	25	52.2	6.21	1.24
Extension Items	**N**	**Mean (%)**	**SD (%)**	**SE (%)**
Overall	65	56.4	11.58	1.44
Aware	40	57.3	12.11	1.92
Unaware	25	55.0	10.76	2.15
Control	25	48.7	9.22	1.84

Table 7. Participants' performances on critical items and extension items

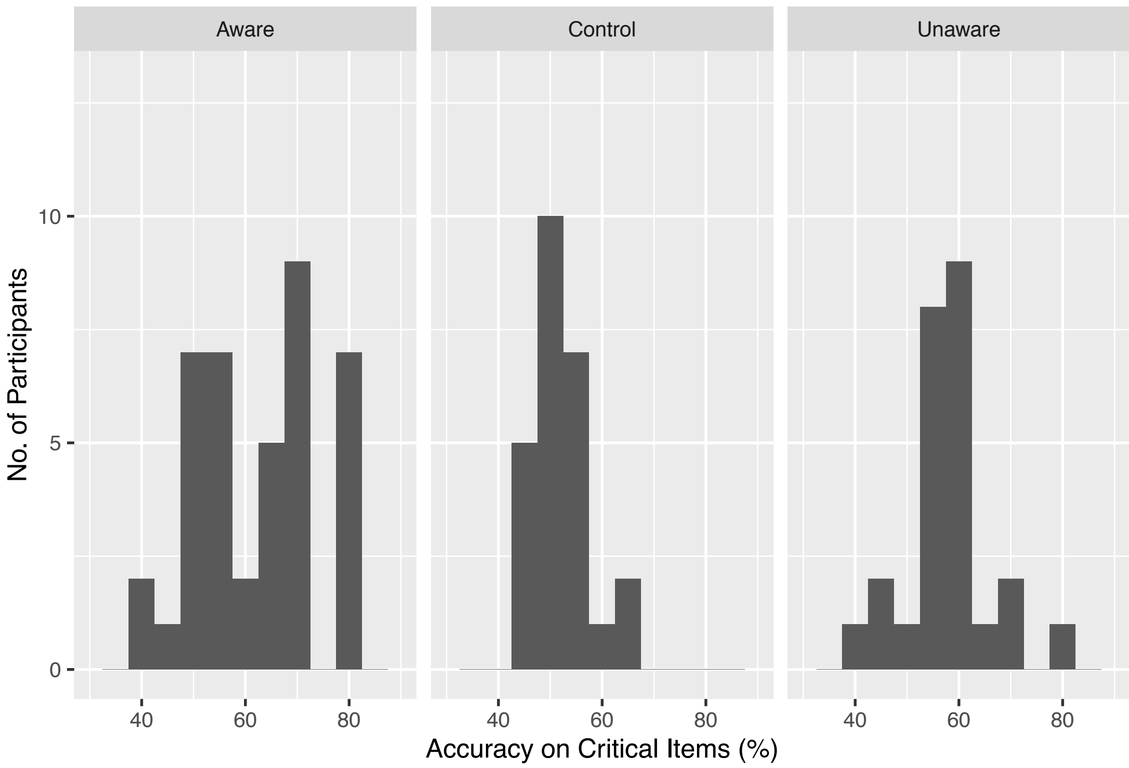

Figure 6. Individual performance on the critical Items by the aware, unaware, and control groups.

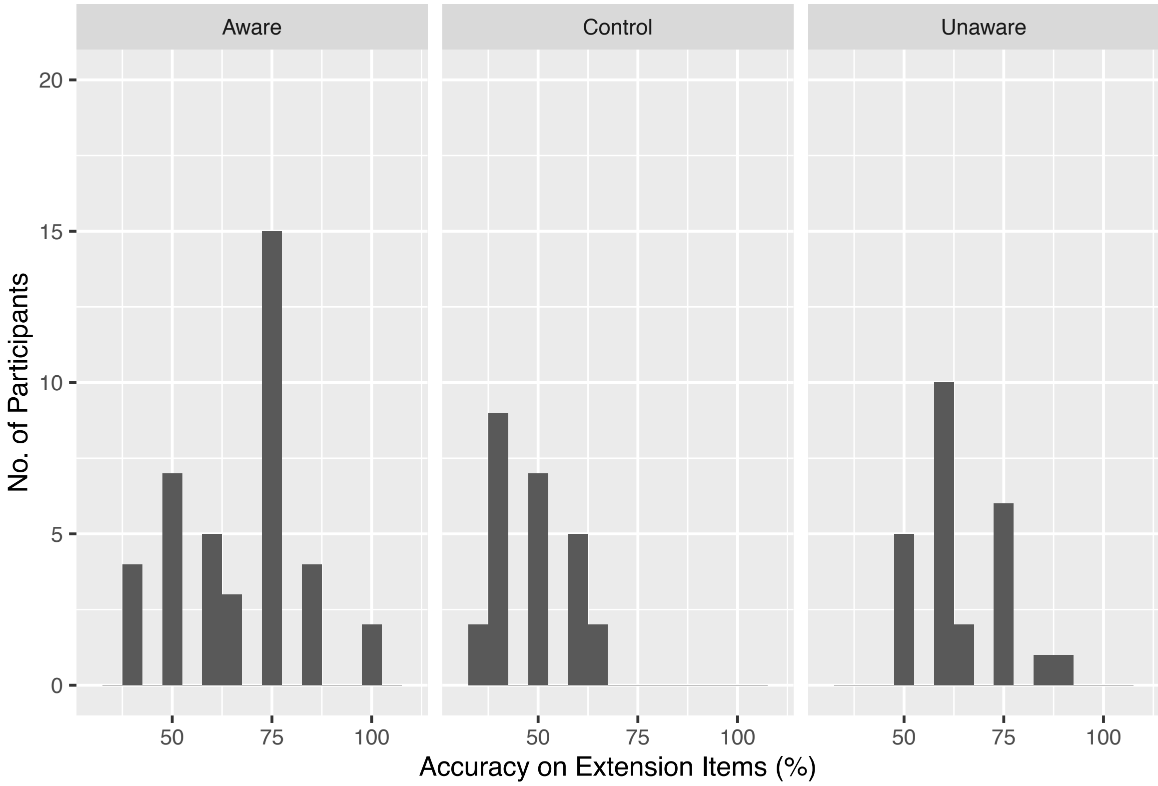

Figure 7. Individual performance on the extension Items by the aware, unaware, and control groups.

Participants performed similarly for the extension items. Overall, the percentage of correct response participants in the whole experimental group on the extension items was significantly above chance, $t(64) = 4.46$, $p < 0.001$, $d = 0.79$, revealing that their knowledge of the target stress patterns was abstract enough to support transfer to novel words with previously unencountered final phonemes. Again, both the aware and the unaware groups performed with significantly above-chance accuracy, $t(39) = 3.81$, $p < 0.001$, $d = 0.86$ and $t(24) = 2.32$, $p = 0.024$, $d = 0.67$ respectively. On the other hand, the control group did not achieve above-chance accuracy, $t(24) = 0.723$, $p = 0.47$, and independent two-tailed t test reveals that the unaware participants performed only marginally significantly better than the control group, $t(48) = 2.24$, $p = 0.030$, $d = 0.65$. Further analysis shows that participants' overall performance on the critical items was significantly better than that on the extension items, $t(64) = 2.39$, $p = 0.018$, $d = 0.422$

(independent two-tailed). This suggests that participants were less accurate when transferring their knowledge of stress patterns to novel words with unheard final phoneme. Also, no significant difference was found on their accuracy on vowel-final words and consonant-final words using independent two-tailed t test, $t(64) = 1.37$, $p = 0.17$ for critical items and $t(64) = 1.34$, $p = 0.18$, indicating that participants had mastered the two target rules similarly well.

The guessing criterion stipulates that knowledge is unconscious when participants are guessing but are performing above chance. When participants in the unaware group chose “guess” in the confidence ratings, their percentages of correct response for both critical items and extension items were still significantly above chance: 59.6%, $t(24) = 4.04$, $p < 0.001$, $d = 1.17$; and 63.0%, $t(24) = 5.74$, $p < 0.001$, $d = 1.66$ respectively, indicating that they possessed unconscious judgment knowledge. As for their structural knowledge, “guess” and “intuition” attributions in the source attribution task indicate unconscious structural knowledge, whereas “memory” and “rule” attributions indicate conscious structural knowledge. For both critical and extension items, correct responses of the combined “guess” and “intuition” attributions were significantly above chance, 60.7%, $t(24) = 5.32$, $p < 0.001$, $d = 1.54$ for critical items and 57.1%, $t(24) = 2.88$, $p < 0.01$, $d = 0.83$ for extension items, showing that the unaware participants also possessed unconscious structural knowledge.

DISCUSSION

The present study set out to investigate incidental learning of novel lexical stress rules, with the goal of exploring whether the resultant knowledge can be implicit, abstract and rule-based. In general, the results demonstrated that the experimental group could develop knowledge of novel stress patterns after merely a little amount of incidental exposure without feedback. They performed similarly well for consonant-final and vowel-final words of both critical and extension items, suggesting that the rules for initial stress and final stress were acquired similarly well. Some participants in the experimental group possessed implicit knowledge of the target stress rules, as assessed by a combination of subjective and objective awareness measures. We conclude that implicit knowledge of stress patterns may be acquired through incidental exposure.

One might find it surprising that the experimental group could develop knowledge of the target stress patterns with only brief, incidental exposure. In this regard, participants' prior linguistic knowledge may have contributed to their learning. Their tone language background and familiarity of lexical stress in English might have facilitated their perception of stress (e.g. the use of various acoustic cues for stress perception). Specifically, whilst there is a preference for non-final stress in English as English feet are trochaic with extrametrical final syllables (Hammond, 1999; Hayes, 1995), English also has a quantity-sensitive lexical stress system (Hayes, 1995). It is possible that participants transferred one English lexical stress rule which is that when a syllable is heavy (e.g. CVC rather than CV), it tends to attract stress. Thus, words with either initial or final stress might have sounded natural to our participants. However, even with the potential facilitation by their relevant prior linguistic knowledge, the learning effects of the unaware participants appear to be limited, ranging from 55% for critical items to 58.4% for extension items. In SLA, a better-than-chance performance is still far from target-like and may not be considered good enough. Hulstijn (2002) argues that successful L2 implicit learning may take an extremely long time, as the L2 has to compete with the resources already taken by L1 (Rohde & Plaut, 1999). In our experiment, since participants only had brief incidental exposure to the target patterns (144 training trials), it is not surprising that they did not exhibit fully target-like behaviour. It would be interesting to test the effects of long-term and richer incidental exposure on the acquisition of implicit phonological knowledge in future research.

A major question in implicit or incidental learning research concerns the nature of implicit knowledge. Specifically, to what extent can implicit knowledge be abstract and rule-based? In our study, the unaware participants were able to apply their knowledge not only to novel words whose endings were encountered in the training phase, but also to novel words with unseen/unheard endings. This reveals that implicit knowledge may be represented at a sufficient level of abstraction to facilitate transfer to completely new lexical items instead on relying solely on their memory of the surface features of the training items. These findings also suggest participants may potentially have acquired rule-based knowledge of the target stress patterns. While our target rules can be described in terms of the mappings between stress assignment and the abstract categories of the final phoneme (i.e. consonant vs. vowel), they might have learnt rules about connection of stress placement with other correlated features such as the length of the word (the number of phonemes) in the word (consonant-final words consist of five phonemes and vowel-final words four phonemes), or with abstract syllable types (the second syllable is closed for consonant-final words, and open for vowel-final words). While our design did not allow us to establish what exactly learners have acquired, in any of the above cases participants possessed abstract and potentially rule-based knowledge of stress rules rather than relying on merely their memory of the surface features of the training items. Still, it is worth noting that participants' performance on extension items was lower than on critical items. This is consistent with previous studies using the transfer technique (e.g. Altmann et al., 1995) that judgment performance was lower on test items with different surface features (e.g. different letter set) or in different modalities (e.g. from visual to auditory).

One might imagine that with the degree of artificiality and simplicity in the experimental design, the present study bears little resemblance to what L2 learners may encounter in real life and it remains unclear to what extent the present findings can be generalized to SLA in naturalistic settings. However, the artificiality and simplicity in our experimental design should not compromise the relevance of the present study to our understanding of the language learning process. Specifically, some nonce words were used and the stimuli were sound files generated by a speech synthesizer instead of recordings by a native speaker. The justification is that to study whether the target stress rules can be learnt, we need to make sure that consonant-final words and vowel-final words do not differ on other features (e.g. the number of syllables; the frequency of the phoneme in other positions) so that participants would not have associated stress location with other correlated features. The inclusion of some nonce words allowed comprehensive control on other confounding variables and avoided the phonological complexity and morphological irregularities found in many languages. In this way, participants could be exposed to systematic data, without which the nature of the participants' knowledge would have been unclear. Also, the use of synthesized speech stimuli, although less natural than recordings by a native speaker, avoided the possibility that participants rely on features found in recordings by a real speaker such as speaker's fluency and within-speaker variation in duration, loudness and pitch to determine stress location in the pronunciation judgment task. Thus, the careful control on the stimuli in this study in fact provides a strong demonstration that participants' above chance performances in the pronunciation judgment task can only be attributed to the learning of the target stress rules. Coupled with the recent finding that success in artificial language learning experiment correlates positively with indices of L2 learning (Ettlinger, 2016), it is not unrealistic to believe that the present findings can help improve our understanding of L2 learning in naturalistic settings.

A factor that was not explored in this study is the potential effects of phrase-final lengthening in stress perception. Phrase-final lengthening generally refers to the lengthening of a rhyme at the end of a prosodic constituent which may serve as a cue for the perception of a phrase boundary (Scott, 1982). In the present study, the duration of the vowels were manipulated as a function of whether the vowels was stressed or unstressed but not their position in the disyllabic words; one might wonder if the effects of phrase-final lengthening might have affected our results. In principle, both stressed and unstressed syllables at phrase-final positions may receive a certain degree of position lengthening. Specifically in the present study, for stress perception based on durational cue in the disyllabic words, the two types of lengthening—stress and phrase-final positions—could be complementary when the stressed syllable is also word-final, or they can compensate for each other when the stressed syllable is word-initial. In both cases our participants might have expected the final syllables to be longer which would lead to a bias towards choosing word-initial stress in the pronunciation judgment task. However, the fact that our participants performed similarly well for words with initial stress and words with final stress suggests that such basis seems to be not non-significant.

In the present study, participants were exposed to only one voice and highly regular patterns. However, in real life SLA settings, most learners are exposed to more than one speaker of the target language, and the input can be highly variable with many exceptions to target patterns. The issues of speaker variability and exceptions in the input in incidental language learning are largely unexplored and warrant further research.

As far as the assessment of awareness is concerned, most previous studies only used one (type of) awareness measure, which may not be adequate for capturing the multifaceted nature of awareness. The present study is one of the few in which participants' awareness of their knowledge was assessed by a combination of subjective and objective measures. The use of more than one awareness measure serves to deal with the complex and multidimensional nature of the phenomenon of (un)consciousness. This ensured that participants in the implicit group possessed implicit knowledge of the target patterns, and demonstrated that acquiring implicit knowledge through incidental exposure is possible. Specifically, retrospective verbal reports tackled participants' verbalisable knowledge; confidence ratings and source attribution assessed the conscious state of participants' judgment knowledge and structural knowledge; and PDP assessed awareness objectively based on participants' behaviour. Although the use of multiple awareness measures contributed to an overall high awareness rate, these awareness measures complemented one another and demonstrated the acquisition of implicit knowledge more convincingly. Given the difference in the nature of these awareness measures and the fact that they are sensitive to different aspects of awareness, their relative effectiveness cannot be compared directly. Still, our data suggest that retrospective verbal reports constitute a less sensitive awareness measure than the other two measures, potentially attributable to the fact that that knowledge becomes verbalisable only when a relatively high level of awareness is reached.

CONCLUSION

The contributions of the present study to the field of language learning are three-fold. First, since the possibility of learning prosodic features in a new language without intention has received relatively little attention, the present study adds to the growing body of evidence on the incidental learning of novel prosodic patterns. Second, our study contributes to the debate on the nature of implicit knowledge by showing that incidental learning may result in abstract and potentially rule-based knowledge. Third, the present study demonstrated the combined use of both subjective and objective awareness measures in an incidental learning research, and provided a strong piece of evidence of the acquisition of implicit knowledge.

REFERENCES

Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 899-912.

Baars, B. J. (2003). How brain reveals mind: Neuoimaging supports the central role of conscious experience. *Journal of Consciousness Studies, 10*(9-10), 100-114.

Bailey, T., Plunkett, K., & Scarpe, E. (1999). A cross-linguistic study in learning prosodic rhythms: Rules, constraints, and similarity. *Language and Speech, 42*(1), 1-38.

Bauer, R. S., & Benedict, P. K. (1997) *Modern Cantonese phonology*. Berlin: Mouton de Gruyter.

Bell, A. (1977). Accent placement and perception of prominence in rhythmic structures. In L. Hyman (Ed.), *Studies in stress and accent: Southern California Occasional Papers in Linguistics, 4* (pp. 1-13). Los Angeles: Dept. of Linguistics, University of Southern California.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*, 316-323.

Chan, C. (1992). *Implicit cognitive processes: Theoretical issues and applications in computer systems design*. Unpublished DPhil thesis, University of Oxford.

Chan, M. K. (2007). *The perception and production of lexical stress by Cantonese speakers of English*. MPhil Dissertation. University of Hong Kong, Hong Kong.

Chan, R. & Leung, J. (2014). Implicit learning of L2 word stress regularities. *Second Language Research*, *30*, 4, 463-483.

Chao, Y. R. (1947). *Cantonese primer.* Cambridge, MA: Harvard University Press.

Chao, Y. R. (1980). Chinese tone and English stress. In: Waugh, L. R. and VanSchooneveld, C. H. (eds) *The melody of language*. Baltimore, MD: University Park Press, 41–44.

Curran, T. (2001). Implicit learning revealed by the method of opposition. *Trends in Cognitive Sciences, 5*, 503-504.

Cutler, A. (1986). Phonological structure in speech recognition. *Phonology Yearbook, 3*, 161-178.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8*(2), 343-350.

Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies, 11*(9), 25-45.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research, 168*, 49-64.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1322-1338.

Dienes, Z., & Berry, D. C. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review, 4*, 3-23.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research-Psychologische Forschung, 69*(5-6), 338-351.

Dulany, D. E., Carlson, R. A., & Deway, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract. *Journal of Experimental Psychology: General, 113*(4), 541-555.

Dunn, J. C., & Kirsner, K. (1989). Implicit memory: Task or process? In S. Lewandowsky, J. C. Dunn & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 17-31). Hillsdale, NJ: Erlbaum.

Dutoit, T., Pagel, N., Pierret, N., Bataille, O., & Vrecken, Van der. (1996). The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *Proceedings of ICSLP'96, Philadelphia, 3*(1393-1396).

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. (2016). The relationship between artificial and second language learning. *Cognitive science*, 822-847.

Face, T. (2005). Syllable Weight and the Perception of Spanish Stress Placement by Second Language Learners. *Journal of Language and Learning*, 36-48.

Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, *36*, 1-34.

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental vocabulary learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, *38*, 44-53.

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*(2), 394-406.

Hammond, M. (1999). *The Phonology of English*. Oxford: Oxford University Press.

Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning? In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365–382). Berlin, Germany: Mouton de Gruyter.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Hualde, J. L. (2005). *The sounds of Spanish*. Cambridge: CUP.

Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193-223.

Hulstijn, J. H. (2005). Theoretical and Empirical Issues in the Study of Implicit and Explicit Second-Language Learning. *Studies in Second Language Acquisition, 27*, 129-140.

Hulstijn, J. H. (2008). 12 Incidental and Intentional Learning. *The handbook of second language acquisition*, *27*, 349.

Hyman, L. (1977). On In L. Hyman (Ed.), *Studies* (Vol. 4, pp. 1-13). Los: Dept. .

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language, 30*, 513-541.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139-154.

Jamieson, R. K., & Hauri, B. (2012). An exemplar model of performance in the artificial grammar task: Holographic representation. *Canadian Journal of Experimental Psychology*, *66*(2), 98–105.

Jamieson, R. K., & Mewhort, D. J. K. (2011). Grammaticality is inferred from global similarity: A reply to Kinder (2010). *Quarterly Journal of Experimental Psychology, 64*, 209-216.

Jiang, S., Zhu, L., Guo, X., Ma, W., Yang, Z., & Dienes, Z. (2012). Unconscious structural knowledge of tonal symmetry: Tang poetry redefines limits of implicit learning. *Consciousness & Cognition*, *21*, 476–486.

Johnstone, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology*, *42*, 61–112.

Kinder, A., & Assmann, A. (2000). Learning artificial grammars: no evidence for the acquisition of rules. *Memory & Cognition*, *28*, 1321–1332.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 79-91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 169-181.

Lai, W. W. S., Wang, D., Yan, N., Chan, V. & Zhang, L. (2011) Influence of English donor word stress on tonal assignment in Cantonese loanwords: An acoustic account. *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong, 1162–65.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*: Routledge.

Leung, J. & Williams, J. (2011). The Implicit Learning of Mappings between Forms and Contextually-Derived Meanings. *Studies in Second Language Acquisition*, *33,* 1, 33-55.

Leung, J. & Williams, J. (2012) Constraints on Implicit Learning of Grammatical Form-Meaning Connections. *Language Learning*, *62*, 2, 634–662.

Leung, J. & Williams, J. (2014) Cross-Linguistic Differences in Implicit Language Learning. *Studies in Second Language Acquisition*, *36*, 4, 733-755.

Luke, K. K. (2000) Phonological re-interpretation: the assignment of Cantonese tones to English words. Unpublished paper presented at the *9th International Conference on Chinese Linguistics*, Singapore.

Morton, J. & Jassen, W. (1965). Acoustic Correlates of Stress. *Language and Speech*, *8*, 3, 159-181.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.

Overgaard, M., Timmermans, B., Sandberg, K., & Cleeremans, A. (2010). Optimizing subjective measures of consciousness. *Consciousness and Cognition, 19*(2), 682-684.

Perruchet, P. (1994). Learning from complex rule-governed environments: On the proper functions of non-conscious and conscious processes. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and non-conscious information processing*: MIT Press.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119*, 264-275.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin, 133*(2), 227-244.

Pothos, E. M., & Bailey, T. M. (2000). The role of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 847-862.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behaviour, 6*, 855-863.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Rebuschat, P. (2008). *Implicit learning of natural language syntax*. Ph.D. thesis, University of Cambridge, Cambridge.

Rebuschat, P. (2013). Measuring Implicit and Explicit Knowledge in Second Language Research. *Language Learning, 63*(3), 595-626.

Rebuschat, P., Hamrick, P., Sachs, R., Riestenberg, K. & Ziegler, N. (2013). Implicit and explicit knowledge of form-meaning connections: evidence from subjective measures of awareness In J. Bergsleithner, S. Frota. & J. K. Yoshioka. (Eds.) *Noticing and second language acquisition: studies in honor of Richard Schmidt*. Manoa: University of Hawaii Press.

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, 33, 829–856.

Rohde, D. L. T., and Plaut, D. C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition*, *72*, 67-109.

Rohrmeier, M, Fu, Q., & Dienes, Z. (2012). Implicit learning of recursive context-free grammars. *PLoS ONE, 7*, e45885.

Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure. *Consciousness & Cognition*, *20*, 214–222.

Rogers, J., Revesz, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional L2 morphology. *Applied Psycholinguistics*.

Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford: Oxford University Press.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Psycholinguistics, 11*, 129-158.

Schmidt, R. (2001). Attention, In P. Robinson (Ed.) *Cognition and second language instruction* (pp. 3-32). Cambridge: CUP.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In. W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan & I. Walker, Proceedings of CLaSIC 2010, Singapore, December 2-4 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Scott, D. R. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America*, *71*(4), 996-1007.

Scott, R., & Dienes, Z. (2010). Knowledge applied to new domain: The unconscious succeeds where the conscious fails. *Consciousness and Cognition, 19*, 391-398.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592–608.

Seth, A. K. (2008). Theories and measures of consciousness develop together. *Consciousness and Cognition, 17*, 986-988.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning system. *Behavioral and Brain Sciences, 17*, 367-447.

Tong, X., Lee, S. M. K., Lee, M. M. L., & Burnham, D. (2015). A tale of two features: Perception of Cantonese lexical tone and English lexical stress in Cantonese-English bilinguals. *PloS one*, *10*(11), e0142896.

Trubetzkoy, N. S. (1969). *Principles of phonology*. Berkeley: University of California Press.

Tunney, R. J. (2005). Sources of confidence judgments in implicit cognition. *Psychonomic bulletin & review*, *12*(2), 367-373.

Tunney, R. J., & Altmann, G. T. M. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35*, 195-202.

Tunney, R. J., & Shanks, D. R. (2003). Does opposition logic provide evidence for conscious and unconscious processes in artificial grammar learning?. *Consciousness and cognition*, *12*(2), 201-218.

Turk, A. & Shattuck-Hufnagel, S. (2007). Phrase-final lengthening in American English. *Journal of Phonetics*, *35*(4), 445-472.

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 195–202.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, 27, 269 – 304.

Williams, J. N. (2009). Implict learning in second language acquisition. In C. R. Williams & T. K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition* (pp. 319-355): Emerald Group Publishing Limited.

Williams, J. N., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. Lingua, 118, 522–553.

Wong, W. Y. P., Chan, M. K., & Beckman, M. E. (2005). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In S. A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing*, *1*, 271-300. Oxford, New York: Oxford University Press.

Zellers, M., Post, B., & Williams, J. N. (2011). Implicit Learning of Lexical Stress Patterns. *Proceedings ICPhS Hong Kong*.

Ziori, E. & Dienes, Z. (2006). Subjective measures of unconscious knowledge of concepts. *Mind and Society*, 5, 105-122

Appendix I:

Duration of the consonants used in the experiment. The same durations were used regardless of the position of the consonants in the words.

Consonant	Duration (ms)	Consonant	Duration (ms)
/b/	65	/d/	65
/f/	70	/g/	65
/k/	65	/ʎ/	80
/m/	80	/n/	80
/p/	65	/s/	70
/t/	65	/ɾ/	80
/r/	80	/l/	80
/θ/	80	/x/	70

[i] The term “stress” has been used in the literature to cover conceptually different senses: 1) relative syllable salience in a string of syllables (also called ‘prominence’); 2) stress in a word as part of the lexical phonology; and 3) stressing of words in an utterance for different propositional and expressive meanings (also called ‘accent’) (Kohler, 2008). While in this paper the term “stress” is used to refer to sense 2), it should be noted that the disyllabic words isolation we used are in principle similar to phonological phrases consisting of an individual word, and the potential effects of pitch accent may be in play.