RESEARCH GRANTS COUNCIL

Application for Allocation from
the Early Career Scheme for 2017/18
Application Form (ECS1)

- **Please read the Explanatory Notes ECS(2) (Aug 16) carefully before completing this form.**
- **To safeguard the interests of the researcher and the university, awardee university bears the primary responsibility for prevention, detection and investigation of research misconduct, including but not limiting to misusing of funds, data falsification, plagiarism and seeking duplicate funding for projects which the applicant has completed substantially or entirely. Concerning research grant applications, the university is strongly advised to use anti-plagiarism software before submitting the application to the RGC.**

PART I SUMMARY OF THE RESEARCH PROPOSAL

[To be completed by the applicant(s)]

1. Particulars of the Project

(a) (i) Name and Academic Affiliation of Principal Investigator:

Name	Post	Unit/ Department/ University
Dr POLITZER-AHLES, Stephen	Assistant Professor	Department of Chinese and Bilingual Studies/The Hong Kong Polytechnic University

(ii) The applicant is within

- ☑ **first year** **(entering into employment between 18 November 2015 to 1 May 2017)**
- ☐ **second year** **(entering into employment between 18 November 2014 to 17 November 2015)**
- ☐ **third year** **(entering into employment between 18 November 2013 to 17 November 2014)**

of his/her full time academic job as an Assistant Professor or career equivalent involving teaching and research duties and the applicant is in substantiation track / tenure track position at the time of submission of the proposal and has not been granted ECS grants or awards.

The applicant confirms that he/she had not been funded under previous Early Career Scheme exercises.

-

(iii) Title of Project: Mental representations of Chinese tones：abstract vs. episodic accounts

(iv) Nature of Application

New ☑ Re-submission ☐ Continuation ☐

(b) (i) Primary Field: Cognitive Neuroscience of Language & Code 4110

Secondary Field: Psycholinguistics & Code 4109

(ii) A maximum of five keywords to characterise the work of your proposal (a maximum of 30 characters for each keyword)

1) neurolinguistics
2) EEG
3) tone sandhi
4) Mandarin
5) lexicon

(iii) Project Duration: 36 Months*

* for longer term projects, please explain in your research plan in Part II 2(b)(i) why the proposed research cannot be completed within the normal span of 36 months.

(iv) Total Amount Requested: $ 938,601

(c) Abstract of Research comprehensible to a non-specialist (a maximum of one A4 page in standard RGC Format for attaching PDF documents or a maximum of 400 words for direct input in the text box):

Attached 1 page(s) as follows.

A major question in the neurolinguistics concerns how words and sounds are stored in the mind, specifically, whether mental representations of words are abstract (containing only a few crucial features) or episodic (containing substantial detail about the person's lifetime experience with that word), or some combination of both. Chinese languages are a fertile testing ground for this issue because systematic tonal variations can lead to differences between a tone's standard canonical form and its context-conditioned variant forms. These differences allow one to test whether or not variant forms leave episodic traces in the brain.

This project will focus on the representation of syllables like *ka*^{R}, which do not exist as canonical forms in the language (i.e., a dictionary would not include any syllables with this pronunciation) but which are commonly encountered (e.g., 卡 has this pronunciation in a sentence like 我電腦卡死了, because of Tone 3 Sandhi which changes *ka*^{L} into *ka*^{R}—where ^{L} indicates a low tone and ^{R} a rising tone). I will use neurolinguistic techniques to reveal how such syllables are represented in the mind. If the mind consists of abstract representations, and variant forms are derived procedurally from them, then such syllables would not have mental representations (since the comprehension system would recognize these syllables as systematic variants that it can derive from other canonical forms) and would not be treated as existing words of the language. On the other hand, if the mind consists of representations with episodic detail, then such syllables may be stored like words (since language users have experience hearing and saying them, just as they do for other words). This project will use neural (Mismatch Negativity and N400) measures in three experimental paradigms that are sensitive to lexicality, to test three complementary hypotheses: (a) such syllables have word-like mental representations; (b) such syllables have nonword-like mental representations; and (c) their mental representations are different from both words and nonwords. These investigations will shed light on how non-canonical but derivable forms are represented in the mental lexicon and, in turn, whether the mental lexicon is arranged abstractly or episodically, particularly in tone languages.

In addition to shedding light on the mechanisms of language comprehension and extending core psycholinguistic theory, the results also have potential applications for the treatment and diagnosis of neurological disorders affecting speech, for second language pedagogy, and for speech technology.

(d) Special funding template

- ☐ **Employment of Relief Teacher under Humanities and Social Sciences Panel (see Enclosure I) (only available for applications under Humanities and Social Sciences Panel)**
- ☐ **Support for Academic Research related to Public Policy Developments (see Enclosure II)**

PART II DETAILS OF THE RESEARCH PROPOSAL

[To be completed by the applicant(s)]

RESEARCH DETAILS

1. Impact and objectives

(a maximum of 800 words in total for the long-term impact and project objectives)

(a) Long-term impact:

How words and sounds are represented and stored in the mind is a major question in the study of language. The same word may be realized in different ways (for example, every word is influenced by coarticulation from neighboring words), but usually we effortlessly recognize words in speech, even when they appear in contexts that we have never heard before. One potential explanation for this ability is that mental representations of language are abstract, not including all the details of how words are pronounced, and the language comprehension system includes mechanisms for matching heard sounds to stored representations while stripping away the phonetic details. Another explanation is that representations of language are episodic, in that the mind remembers many different instances of a given word or sound that it has heard, in order to use this experience with linguistic variability to aid comprehension.

The proposed project will investigate this issue using Chinese tones, which undergo numerous changes in context and thus raise questions of how the different variants of the same word or same tone are represented in the brain and processed during speech comprehension. This project focuses on Mandarin Tone 3 Sandhi, and specifically on syllables like ká, which only exist as pronunciation variants of other syllables, rather than as meaningful citation forms in their own right (as described in the Abstract and the Research Plan). We will measure whether the neural representations of these syllables are word-like or not. Abstractionist theories of lexical representation predict that such syllables are not stored in the brain, because they can be derived from an existing abstract representation (e.g. kǎ), whereas episodic theories of lexical representation predict that these are stored in the brain because speakers hear and say them throughout their lifetimes. Likewise, abstractionist theories predict that the frequency with which derived forms are experienced (e.g., guǎn occurs very often in contexts where it will be pronounced guán [a sandhi allomorph similar to ká], but zhuǎn does not) does not influence their lexical representations, whereas episodic theories predict that it does. Both sets of predictions are testable using neurolinguistic and psycholinguistic paradigms.

In addition to the phonologically conditioned alternation described above, Mandarin has many other alternations (e.g., half-third sandhi and tonal coarticulations) with very different linguistic properties; other Chinese languages have an even wider variety of tone-related alternation phenomena. While the proposed study only examines Mandarin Tone 3 Sandhi, the results will establish a benchmark against which other tonal phenomena can be compared in future research on other tonal alternations in Mandarin and other Chinese languages.

Tone languages are under-represented in research on speech comprehension thus far: most of what we know about the mental representation of language comes from research on Indo-European languages, and aspects of language that work differently than tone.

Therefore, study of how tones are mentally represented is necessary for a broad understanding of how the brain processes a wide variety of languages. Tones still pose a challenge to many core concepts in psycholinguistics (for example, there is disagreement over how to account for tone when calculating the size of phonological neighborhoods—which in turn play a major role in many aspects of language processing—and there are open questions about the relationship between tonal and segmental information in speech and lexical activation). It remains an open and important question whether tones are represented and processed in a similar way as segments. Thus, the proposed research on how tonal variation is mentally represented could have major implications for psycholinguistic theory.

Finally, while the primary goal of this project is to inform our understanding of the basic science underlying linguistic representation and processing, this line of research also stands to have broader, translational impacts. There are many applications for knowledge about how speech is represented and comprehended. Many neurological conditions selectively impact certain aspects of speech and language comprehension, and better understanding the brain mechanisms that subserve these processes may eventually inform the treatment or diagnosis of such conditions. Likewise, understanding how sounds are represented in the mind of native speakers can improve second-language pedagogy. Finally, speech recognition software faces the same challenges that human listeners do, and thus learning how the human language architecture handles this problem helps inform computational solutions.

(b) Objectives

[Please list the objectives in point form]

1. To test whether lexical representations of tone are abstract or episodic
   - Experiments 1a & 1b: Using Mismatch Negativity to examine if sandhi accidental gaps [see Background of Research] behave like nonwords (suggesting abstract representation) or words (suggesting episodic representation)
   - Experiments 2a, 2b, & 3: Using Mismatch Negativity and N400 to test whether allomorph surface frequency is lexically represented or not
2. To extend theories of Indo-European languages' mental representation into Chinese tones
3. To establish benchmarks for how these phenomena work in Mandarin, informing future research in other tone systems

2. Background of research, research plan and methodology:
(a maximum of seven A4 pages in total in standard RGC format for items (a) and (b)(i); a maximum of one A4 page for item (b)(ii))
(a) Background of research
(b) (i) Research plan and methodology

THE MENTAL REPRESENTATION OF CHANGING SOUNDS

Sounds in natural language vary: a given sound may be realized in different ways depending on the context. For example, the phoneme /t/ in many English dialects may be pronounced as [t] (*s**t**op*), [t^{h}] (***t**op*), [t̚] (*ca**t***), [ɾ] (*wa**t**er*), or [ʔ] (*bu**tt**on*). In tone languages, supra-segmental aspects of sound show similar variation. For example, Mandarin exhibits *tone sandhis* (systematic patterns of context-conditioned tone change) such as Tone 3 Sandhi, whereby an underlyingly Tone 3 syllable can be pronounced as Tone 2 in an appropriate context: e.g., 卡 is typically pronounced *ka*^{L} (with low tone), but is instead pronounced *ka*^{R} (with rising tone) in a context such as (1), where it is followed by another Tone 3 within the same intonational unit:

(1)	我	電腦	卡	死了！
Underlying form:	*wo*^{L}	*dian*^{F}*nao*^{L}	***ka***$^{\mathbf{L}}$	*si*^{L}*-le*
Surface form:	*wo*^{L}	*dian*^{F}*nao*^{L}	***ka***$^{\mathbf{R}}$	*si*^{L}*-le*
Gloss:	my	computer	slow	Intensifier
Translation:	"My computer is slow as heck!"			

A major question in psycholinguistics concerns how the mind's language architecture handles such variation. Abstractionist models of the lexicon (e.g., Coltheart, 1980) argue that the lexical representation of a syllable (i.e., the way that a syllable is stored in the mind) is composed of its abstract citation form plus knowledge of morpho-phonological rules or constraints that can derive different surface forms from that citation form. For example, for a syllable like 卡, the brain would only store its canonical pronunciation (*ka*^{L}) and knowledge about Tone 3 Sandhi; this knowledge can be used to realize other pronunciations like *ka*^{R}. On the other hand, episodic, full models of the lexicon (e.g., Goldinger, 1998) argue that lexical representations are composed of all possible surface realizations of that syllable, stored together in the lexicon. For example, because speakers of Mandarin hear 卡 pronounced as both *ka*^{L} and *ka*^{R} in their daily lives, both representations (i.e., both the underlying form and the surface form) are stored alongside one another in the mind, with no need for abstract phonological rules (rather, such rules may arise as an epiphenomenon of the episodic representations). This project tests these accounts by measuring brain activity during the comprehension of Mandarin Chinese syllables which vary because of tone sandhi. In particular, I will take advantage of two linguistic and psycholinguistic properties that shed light on lexical representation: *accidental gaps* and *lexical frequency*.

ACCIDENTAL GAPS AND LEXICAL FREQUENCY IN TONE SANDHI

The set of syllables that exist in Mandarin includes many gaps: not every possible syllable occurs with every possible tone. For example, while *shi* may occur with any tone (*shi*^{H}: 濕 "wet"; *shi*^{R}: 時 "time"; *shi*^{L}: 使 "make"; *shi*^{F}: 是 "is"), *pa* may not: *pa*^{H}, *pa*^{R}, and *pa*^{F} are syllables which correspond to existing morphemes in standard Mandarin, but there is no meaningful syllable *pa*^{L}. These are called *accidental* gaps because there is no principled phonotactic reason why *pa* (for instance) could not occur in Tone 3; just like, e.g., *blick* in English, these syllables are phonotactically legal.

Such gaps may interact with tone sandhi. For example, *ka*^{R} is an accidental gap: there is no Mandarin syllable with a citation form of *ka*^{R}. However, *ka*^{L} is an existing syllable (卡 "blocked/slow"), and sandhi can cause *ka*^{L} to be pronounced as *ka*^{R}, as shown in (1) above.

In short, even though *ka*^{R} is not an existing syllable in Standard Mandarin and would not appear in any dictionary, speakers of the language do hear it: it is a variant of an existing syllable. This makes it different than other accidental gaps, like *pa*^{L}, which are un-derivable (no phonological alternation generates *pa*^{L} from another syllable). A sandhi accidental gap like *ka*^{R}, therefore, is a useful test case for contrasting models of lexical representation. An abstractionist model which assumes that only underlying forms are stored would predict such a syllable to be treated as a nonword by the language architecture, since it is derivable by rule and does not need to be stored in the lexicon. On the other hand, an episodic model would predict such a syllable to be treated as a real word, since it is encountered in the language and thus could be stored in the lexicon. It is also possible that different listeners represent the syllable in different ways. The present study will use neurolinguistic and psycholinguistic methods that are especially suited to testing whether a given syllable behaves like a word or a nonword.

Tone sandhi may also interact with lexical frequency (how often a word is used in the language). While all Tone 3 syllables have the potential to undergo tone sandhi, they differ in how often they do it. For instance, both *guan*^{R} and *zhuan*^{R} are sandhi accidental gaps like *ka*^{R} (i.e., neither exists as a citation form, but *guan*^{L} and *zhuan*^{L} both do), but *guan*^{R} is a high-frequency accidental gap (between 43.8% and 51.3% of underlyingly *guan*^{L} tokens in the Lancaster Corpus of Mandarin Chinese [McEnery & Xiao, 2004] occur in contexts where they will undergo sandhi and be pronounced *guan*^{R}; a large portion of these are the morpheme 管 occurring in contexts 管理 *guan*^{L}*li*^{L}, 管我 *guan*^{L} *wo*^{L}, or 管你 *guan*^{L} *ni*^{L}) whereas *zhuan*^{R} is a low-frequency variant (only between 2.8% and 5.3% of underlyingly *zhuan*^{L} tokens occur in sandhi-triggering contexts). Given that episodic models of the lexicon predict lexical representations to be directly related to the frequency with which a word is experienced, such differences in frequency provide another window into lexical representation. Specifically, episodic models would predict the differences in surface frequency between gaps like *guan*^{R} and *zhuan*^{R} to be reflected in psycholinguistic and neurolinguistic measures of language processing, whereas abstract models assume that these surface forms are not lexically represented and thus such frequency differences should have no consequences for language processing.

HOW TO TEST LEXICALITY AND FREQUENCY EFFECTS

Brain potentials can shed light on the lexical status of a sound. In particular, the *Mismatch Negativity* (MMN) component of the event-related potential is modulated by lexicality. MMN is measured by infrequently presenting one sound (the *deviant*) mixed in with a different sound that is presented much more frequently (the *standard*); e.g., in a continuous series such as *s s s s s s* ***d*** *s s s* ***d*** *s s s s s s s* ***d*** *s s s s*..., the *s* tokens are the standards and the ***d*** tokens the deviants. The brain response to the infrequently presented sound minus the brain response to the frequently presented sound yields the MMN (Näätänen et al., 2007). Across a variety of languages (including Mandarin; Gu et al., 2012) and paradigms, it has been found that real-word deviants elicit higher-amplitude (i.e., more negative) MMNs than nonword deviants (see review in Pulvermüller & Shtyrov, 2006). Pulvermüller and Shtyrov (2006) and Pulvermüller (1999) suggest that this effect may be due to the development of Hebbian cell assemblies: words that are experienced often by a listener become associated with distributed cortical representations, which are strongly activated when that word is perceived, whereas sound patterns that are infrequently experienced and/or not associated with lexical representations do not develop such cortical representations. The MMN, therefore, offers a means to examine whether sandhi accidental gaps like *ka*^{R}, which do not exist as citation forms in the language but which are experienced by listeners because they are allomorphs of existing words, are or are not stored in the lexicon—and, in turn, whether the mental lexicon is organized abstractly or episodically. This same MMN paradigm has also been used to reveal differences between high-frequency and low-frequency words, mimicking the differences between words and nonwords: high-frequency words elicit larger (more negative) MMNs than low-frequency words (Alexandrov et al., 2011). Thus, this paradigm is useful to test both the lexicality and the frequency aspects of sandhi accidental gaps described above.

Finally, because the MMN is often robust in individual subjects, it is an ideal measure to reveal potential individual differences in how listeners represent accidental gaps (e.g., if some listeners represent syllables abstractly but others represent them episodically).

WORK DONE BY OTHERS

As the abstract vs. episodic question is one of the central debates in psycholinguistics, there is a large body of work examining these models of storage in mostly Indo-European languages; while a comprehensive summary of this literature is beyond the scope of this proposal, reviews are available (e.g., Kouider & Dupoux, 2009; Pisoni & Levi, 2007). Across a wide variety of paradigms, evidence has been found for both kinds of models, leading some (e.g., Kouider & Dupoux, 2009) to suggest hybrid models in which both factors play a role. Below I focus on research examining lexical storage of syllables undergoing tone sandhi.

Wiener and Turnbull (2013) found that sandhi gaps were responded to more slowly than other word types in a lexical decision task, suggesting that they caused processing difficulty. They did not, however, directly test whether such syllables are stored in the mental lexicon.

Results regarding the representation of such syllables are mixed. Chien and colleagues (2016) found that a sandhi-undergoing target word was primed by a Tone 3 syllable more than a Tone 2 syllable in auditory priming, suggesting that sandhi syllables are lexically represented in terms of their abstract underlying forms rather than episodically stored surface forms. Regarding speech production, studies using implicit priming (X. Chen, 2012; see also Politzer-Ahles & Zhang, in press) and brain potential recording during word naming (Zhang et al., 2015) found evidence that producing sandhi-undergoing compounds requires first retrieving the Tone 3 underlying form and subsequently re-encoding it into a sandhi form. Such findings could be taken as evidence that lexical representations abstract away from sandhi forms. On the other hand, other studies using picture-word interference (Nixon et al., 2014), mismatch negativity (Li & Chen, 2015), and again implicit priming (Y. Chen et al., 2011) have suggested that both underlying and sandhi variants of Tone 3 syllables may be mentally represented, and directly accessed, in the lexicon. Some studies using the same or similar paradigms as those discussed above, i.e., implicit priming (J. Chen & Chen, in press) and cross-modal priming (Zhou & Marslen-Wilson, 1997), have been inconclusive, failing to provide evidence for either abstract or episodic accounts. Overall, the evidence regarding the lexical representation of sandhi variants remains mixed, with different conclusions suggested by different paradigms, or even different studies in the same paradigm with only slight methodological differences. No studies yet have taken advantage of sandhi accidental gaps and paradigms specifically suited to measuring lexicality to address this question.

As for lexical frequency effects, Wiener and Ito (2015) show behavioural and eye-tracking evidence that syllable-specific tonal lexical frequency plays a role in spoken language comprehension. While this study did not examine tone sandhi, the results pave the way for the current study to fruitfully examine frequency effects in sandhi gaps as described above.

WORK DONE BY THE PI

Pilot work using the MMN lexicality paradigm described above (Politzer-Ahles, ms.; see Figure 1) found preliminary evidence that sandhi accidental gaps like *ka*^{R} may have behaved like nonwords, suggesting that lexical representations may be abstract. This study, however, had only a small number of participants and thus lacked power to make meaningful statistical inferences. A series of MMN studies using a different MMN paradigm (Politzer-Ahles et al., 2016) found evidence that Tone 3 has a phonologically underspecified representation at the neural level (contra Li & Chen, 2015). This study, however, was not specifically about tone sandhi, and thus it is not possible to infer the neural representation of the *sandhi forms* of the syllables that were tested.

Politzer-Ahles & Zhang (in press) used implicit priming of sandhi production and found evidence that the underlying forms of sandhi-undergoing compound words are represented in the mental lexicon. However, the results did not rule out the possibility that surface forms are also represented; furthermore, several conflicting results have also been found (see above), and this result has also failed to be replicated in a similar paradigm (Politzer-Ahles & Connell, submitted).

Outside the topic of Mandarin tone representations, the PI is highly experienced with electrophysiological research in general, having multiple publications and presentations in this area (e.g., Bovolenta, Politzer-Ahles, & Husband, 2016; Fiorentino, Politzer-Ahles, et al., 2015; Hunt, Politzer-Ahles, et al., 2013; Politzer-Ahles, 2015; Politzer-Ahles et al., 2013, 2016; Politzer-Ahles & Gwilliams, 2015; Schluter, Politzer-Ahles, & Almeida, 2016; Tucker, Politzer-Ahles, et al., 2014) and extensive involvement in the EEGLAB and Fieldtrip open-source software communities for EEG data analysis. Thus, between this experience and being situated in a department with an EEG laboratory and collaborators with experience in Mandarin tone research, the PI is well equipped to carry out this line of research.

MOTIVATION

As described above, the findings regarding storage of tone sandhi forms in the mind are currently rather mixed. These findings come from a wide variety of paradigms, some of which are better suited to address this research question than others: for instance, the implicit priming paradigm sheds light on representations that the mind is using as it prepares to articulate speech, which may or may not be the same as representations that are stored in the mental lexicon; and for auditory and cross-modal priming, the mechanisms under investigation are not always clear (for example, experimenters using the same paradigm have attributed opposite predictions to the same theories, depending on whether they expect prime-target match to engender facilitation or competition). Finally, no research as yet has directly tested these accounts for tone sandhi at the neural level (Li & Chen, 2015, and Politzer-Ahles et al., in press, examined Tone 3 in general rather than tone sandhi per se, and Zhang et al., 2015, focused on articulatory preparation rather than representation). The present study, on the other hand, will use a paradigm that is particularly well-suited to test lexical representations, that makes straightforward predictions and allows for a high degree of experimental control, and that examines lexical representation at the neural level.

RESEARCH PLAN AND METHODOLOGY

The present proposal encompasses three sets of experiments. Experiments 1a and 1b will use the MMN paradigm described above (and piloted in Politzer-Ahles, ms.) to examine whether sandhi accidental gaps like *ka*^{R} are stored in the mind as words or not, in order to shed light on whether the mental lexicon preserves episodic representations or only abstract representations.

Experiments 2a and 2b will use a similar paradigm to test the closely related question of whether the frequency of tone sandhi influences the lexical representation of sandhi accidental gaps. Because abstract and episodic accounts of the mental lexicon make divergent predictions about whether sandhi accidental gaps are represented in the lexicon, they likewise make divergent predictions about whether a word's frequency of occurrence in its sandhi form, as opposed to its underlying form, influences the lexical representation. Each of these series of experiments includes a replication to ensure that results from one MMN experiment, which uses only a small set of items, are generalizable to more items (e.g., Experiment 1b is a replication of Experiment 1a with new items, and likewise for Experiments 2b and 2a). Finally, Experiment 3 again tests the role of sandhi surface frequency, using a different experimental paradigm which allows for testing a much larger number of items at once, in order to further examine the robustness of the results across a larger set of words in the language.

Each experiment will be pre-registered through the Open Science Foundation (http://osf.io) prior to the commencement of data collection.

Experiments 1a-b: Testing the lexicality of sandhi accidental gaps using MMN (10 months)

Aim: To examine whether sandhi accidental gaps are or are not stored in the mental lexicon, using a method which directly probes their lexical representation at the neural level.

Design and predictions: 48 native Mandarin speakers from the north of China will be tested in the EEG laboratory in the Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University. (Power analysis [https://politzerahles.shinyapps.io/ERPpowersim/] suggests that N=48 will give 80% power to detect a −1 µV effect based on pilot data from Politzer-Ahles, ms.) While watching a silent movie, participants will hear four blocks of sounds each presented in an MMN oddball design (the repeated syllables represent the standards, and the syllables in bold represent the deviants; see also Table 1):

pai^{R} *pai*^{R} *pai*^{R} ***pa***$^{\mathbf{R}}$… (word deviant)
pai^{L} *pai*^{L} *pai*^{L} ***pa***$^{\mathbf{L}}$… (nonword deviant)
kai^{L} *kai*^{L} *kai*^{L} ***ka***$^{\mathbf{L}}$… (word deviant)
kai^{R} *kai*^{R} *kai*^{R} ***ka***$^{\mathbf{R}}$… (sandhi accidental gap deviant)

Two blocks will contrast unambiguous word (*pa*^{R}) and nonword (*pa*^{L}) deviants. These serve as a manipulation check to confirm that the design succeeded; if the paradigm is sensitive to lexicality, the unambiguous word deviant *pa*^{R} should elicit more a negative MMN than the unambiguous nonword deviant *pa*^{L}. The remaining two blocks contrast the word deviant *ka*^{L} with the sandhi accidental gap deviant *ka*^{R}. If lexical representations are abstract, *ka*^{R} should behave like a nonword, eliciting a less negative MMN than *ka*^{L}, leading to a main effect of lexicality. On the other hand, if lexical representations are episodic, then *ka*^{R} should elicit a similar MMN as *ka*^{L}, leading to an interaction (such that there is a lexicality effect in *pa* blocks but not in *ka* blocks).

While the present study uses syllables in isolation (i.e., not presented in a sandhi-triggering context), the MMN paradigm used still yields different predictions for episodic and abstract accounts of lexical representation. Crucially, the reason that words are argued to yield greater MMNs than nonwords (Pulvermüller, 1999) is because sound-meaning pairings that are frequently experienced are thought to encourage the formation of cortical cell assemblies which fire action potentials in synchrony, and hearing the sound again triggers the cell assembly to fire.

This mechanism is independent of phonological context and thus it is expected that hearing a previously-experienced meaningful syllable would activate cortical representations associated with it, even if it is not occurring in the sort of context where it was experienced before. Thus, lexicality effects are predicted to emerge even without a sandhi-licensing phonological context.

Each block contrasts standards and deviants with different vowels (*ai* vs. *a*), following the design of Shtyrov and Pulvermüller (2002). This is done in order to avoid contrasting different tones within a block (e.g.,, a design with *ka*^{R} standards and *ka*^{L} deviants contrasted in the same block), given that previous MMN research (Li & Chen, 2015; Politzer-Ahles et al., in press) has demonstrated that different tonal contrasts elicit greatly different MMN effects. Also note that in the present paradigm, only the lexicality of the deviants matters; the lexicality of the standards has been shown to not modulate MMN amplitude (Shtyrov & Pulvermüller., 2002). Each block will consist of 180 deviants and 1305 standards, with block order counterbalanced across participants. EEG analysis will use the parameters from in Politzer-Ahles et al. (2016).

Stimuli will be carefully constructed to avoid introducing artifactual modulations of the MMN, which is highly sensitive to acoustic properties of the stimuli. Recordings will be modified from natural utterances using Praat (Boersma & Weenink, 2016), as described in Politzer-Ahles (ms.), to ensure the relevant stimuli differ only in F0 contour.

Experiment 1b will be a replication of Experiment 1a with different stimuli, but with all other parameters held the same. Because this MMN paradigm tests only a small number of items, unlike other psycholinguistic designs which allow for many items, it is important to re-test such an MMN design on a new set of items to increase confidence that the effect is generalizable.

Experiments 2a-b: Testing sandhi gap frequency effects using MMN (10 months)

Even if a sandhi accidental gap such as *ka*^{R} is episodically stored in the lexicon, it may have a lower lexical frequency than *ka*^{L} and thus yield a smaller MMN effect even though it has a lexical representation. In other words, the results of Experiment 1 could turn out to be consistent with either a lexicality effect or a frequency effect. Therefore, Experiment 2 will directly test whether the lexical frequency of a sandhi form, as opposed to an underlying form, influences MMN amplitude, to provide converging evidence with the results of Experiment 1.

Design and predictions: 48 participants will be sampled from the same population as in Experiment 1. Real words with low and high lexical frequencies will be chosen from the Da (2004) corpus. The words given here are an example (see also Table 2); a different set of stimuli may be selected during experiment development. The time allocated for this experiment includes time for corpus processing to develop a more accurate index of sandhi frequency than that available from the Lancaster corpus (McEnery & Xiao, 2004).

zhe^{R} *zhe*^{R} *zhe*^{R} ***ge***^{R}... (high lexical frequency word deviant)
ge^{R} *ge*^{R} *ge*^{R} ***zhe***$^{\mathbf{R}}$... (low lexical frequency word deviant)
zhuan^{R} *zhuan*^{R} *zhuan*^{R} ***guan***$^{\mathbf{R}}$... (high surface frequency sandhi gap deviant)
guan^{R} *guan*^{R} *guan*^{R} ***zhuan***$^{\mathbf{R}}$... (low surface frequency sandhi gap deviant)

As in Experiment 1, the first pair of blocks is a manipulation check: if the experiment successfully detects frequency effects, then the high-frequency *ge*^{R} deviant should elicit a more negative MMN than the low-frequency *zhe*^{R} deviant. The second pair of blocks tests the role of surface frequency of sandhi accidental gaps; crucially, the phonological contrast (*zh* vs. *g* in this example) is exactly the same in both pairs of blocks. If lexical representations are episodic, then the high surface frequency *guan*^{R} deviant should elicit a more negative MMN than the low surface frequency *zhuan*^{R} deviant, leading to a main effect of frequency across the word and sandhi-gap blocks. On the other hand, if lexical representations are abstract, then there should be no MMN difference between *guan*^{R} *and zhuan*^{R}, leading to an interaction (a lexicality effect in the word blocks but not the sandhi accidental gap blocks). Note that, while in Experiment 1 a null interaction is expected under the abstract account and a significant interaction under the episodic account, the predictions are reversed in Experiment 2; this increases the likelihood of finding positive results across the experiments.

In Experiment 2, MMNs will be computed by taking the brain response to a given deviant syllable, minus the brain response to the same syllable when it is presented as a standard in the other block. This subtraction method allows for subtracting out most low-level differences between stimuli (but see Schröger & Wolff, 1996) to minimize the possibility that MMN differences between, e.g., *zhe*^{R} and *ge*^{R} are based on low-level acoustic factors rather than on lexical frequency. This was not possible in Experiment 1, but is possible in Experiment 2, which contrasts high- and low-frequency syllables with the same tone (rather than contrasting words and sandhi, gaps which necessarily have different tones). Other experimental procedures will be the same as Experiment 1.

As in Experiment 1, Experiment 2 will include a replication (Experiment 2b) with a new set of stimuli, and all other parameters held the same as in Experiment 2a.

Experiment 3: Surface sandhi frequency effects on a large sample of words (16 months)

Aims: As noted above, while the MMN paradigm allows for excellent experimental control and yields straightforward predictions for abstract versus episodic models, a limitation is that only a small set of items can be tested per experiment. Thus, Experiment 3 will seek to provide cross-paradigm converging evidence by using a different neurolinguistic paradigm that allows for testing a large sample of words with advanced statistical modeling techniques.

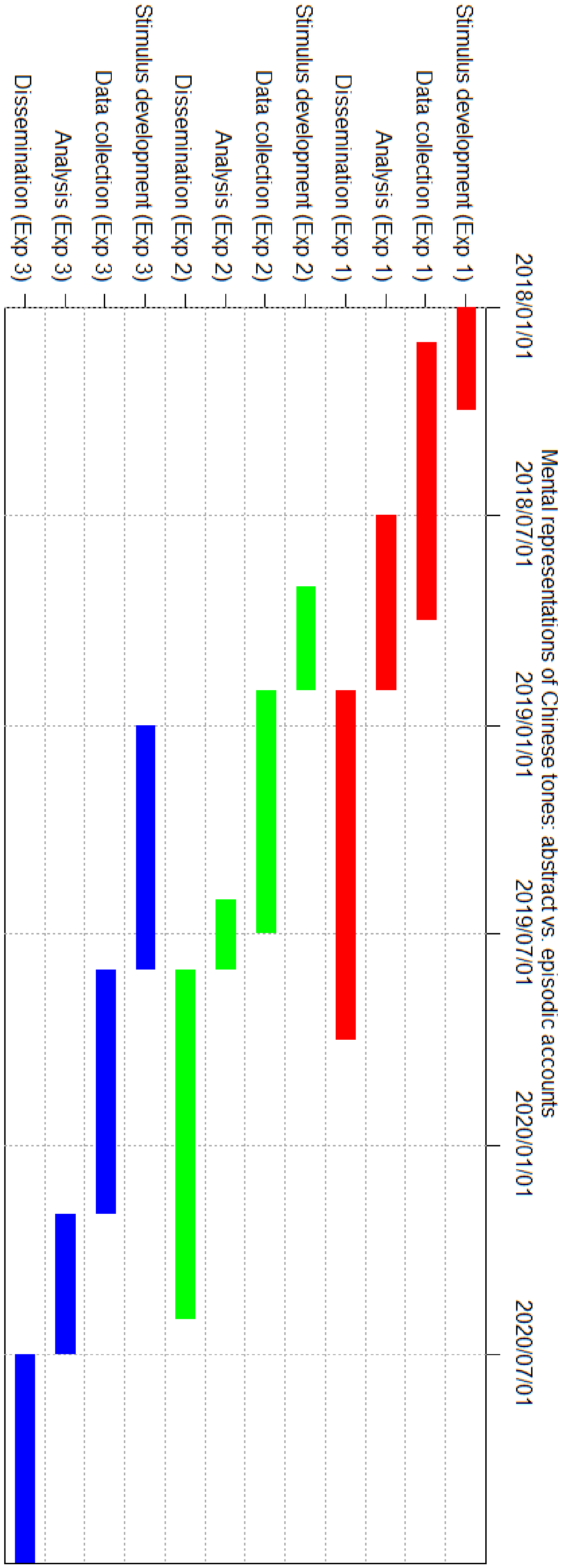
Design: This experiment will use auditory lexical decision, in which the participant hears an isolated syllable and decides whether it is an existing syllable in Mandarin (e.g. Wiener & Turnbull, 2013). In such a paradigm, higher-frequency morphemes are known to elicit smaller (more positive) N400 components in the EEG (Kutas & Federmeier, 2000).

This experiment will include a large number of sandhi accidental gaps in addition to real words and nonwords. Single-trial ERP regression (Hauk et al., 2006; Smith & Kutas, 2015) based on linear mixed-effects models (Baayen et al., 2008) will be used to examine the role of surface frequency for sandhi accidental gaps and their Tone 3 correspondents. If syllables are represented episodically, then the frequency with which a sandhi accidental gap appears in its sandhi form should modulate the amplitude of the N400 and should be independent of the frequency of the syllable's corresponding Tone 3 form. On the other hand, if sandhi gaps are represented abstractly, then their Tone 2 and Tone 3 surface frequencies should not independently modulate N400 amplitude; rather, they should both be associated with the same lexical representation and have the same effect on N400. Log-likelihood tests will be used to statistically evaluate whether surface frequency or shared frequency better accounts for the N400 amplitudes elicited by these syllables.

The experiment will use all Mandarin sandhi accidental gaps. The number of participants will be determined based on the results of power analysis using pilot data from an unrelated N400 experiment currently underway. Because this experiment will require a larger sample of stimuli to prepare and may require more participants, a longer time period has been allocated.

(b) (ii) A one-page Gantt Chart showing the research activities

Attached 1 pages(s) as follows


Mental representations of Chinese tones: abstract vs. episodic accounts
2018/01/01
2018/07/01
2019/01/01
2019/07/01
2020/01/01
2020/07/01
Stimulus development (Exp 1)
Data collection (Exp 1)
Analysis (Exp 1)
Dissemination (Exp 1)
Stimulus development (Exp 2)
Data collection (Exp 2)
Analysis (Exp 2)
Dissemination (Exp 2)
Stimulus development (Exp 3)
Data collection (Exp 3)
Analysis (Exp 3)
Dissemination (Exp 3)

(c) A maximum of two non-text pages of attached diagrams, photos, charts and table etc, if any.

Attached 2 pages(s) as follows

(a)

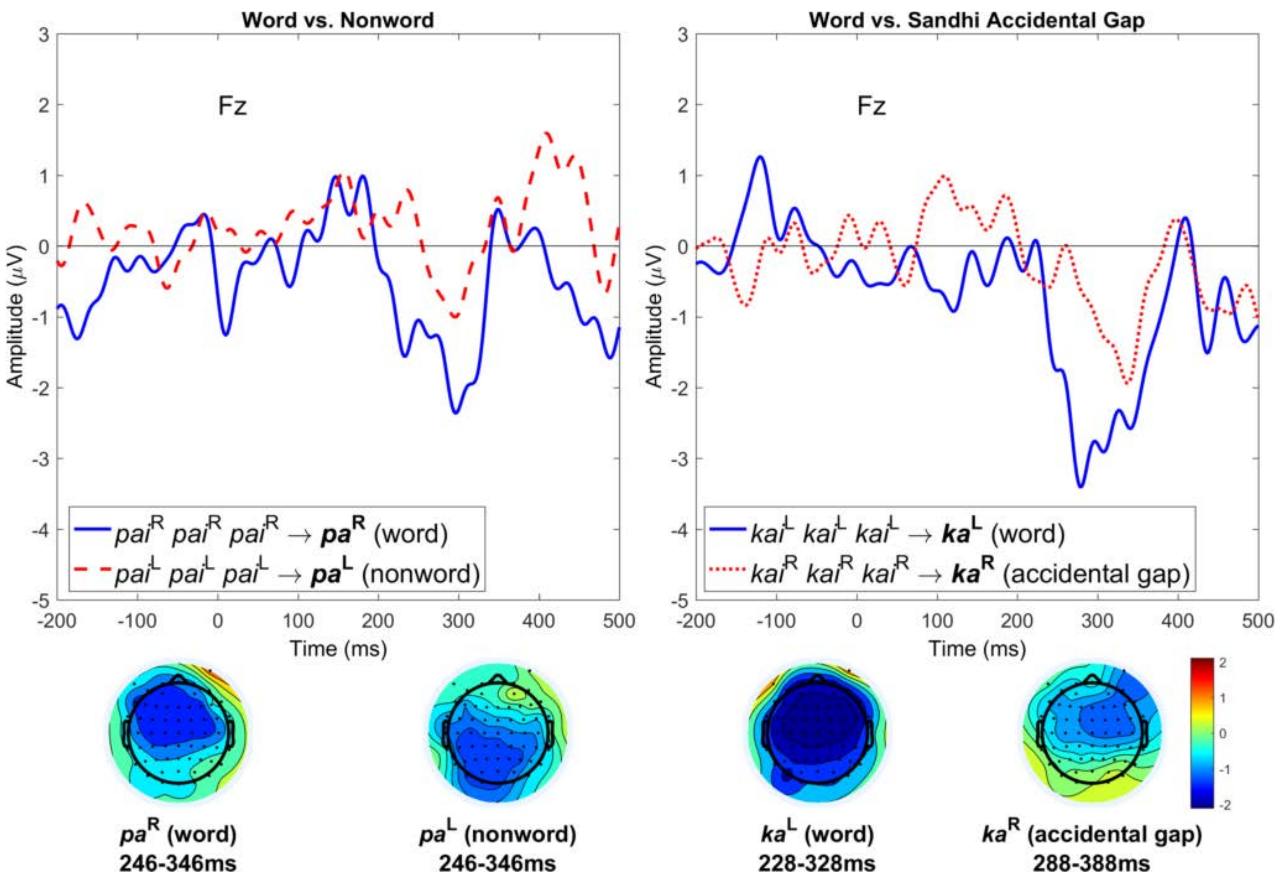


Figure 1. Results from pilot study (Politzer-Ahles, ms.), showing results for the unambiguous word vs. nonword contrast on the left and the word vs. sandhi allomorph contrast on the right.

	Word	**Nonword**
<u>Control contrast</u>: unambiguous word vs. unambiguous nonword	*pa*^{R}	*pa*^{L}
<u>Experimental contrast</u>: unambiguous word vs. sandhi accidental gap	*ka*^{L}	*ka*^{R}

Table 1. 2×2 design of Experiments 1a and 1b.

	High-frequency	**Low-frequency**
<u>Control contrast</u>: frequency in existing words	*ge*^{R}	*zhe*^{R}
<u>Experimental contrast</u>: frequency in sandhi accidental gaps	*guan*^{R}	*zhuan*^{R}

Table 2. 2×2 design of Experiments 2a and 2b.

(d) Reference (a maximum of three pages for references is allowed for listing the publications cited in Section 1-2. All full references should be provided, including all authors of each reference.)

Alexandrov, A., Boricheva, D., Pulvermüller, F., & Shtyrov, Y. (2011). Strength of word-specific neural memory traces assessed electrophysiologically. PLoS ONE, 6, e22999.

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390-412.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer, Version 5.4. [Computer program]. Retrieved from http://www.praat.org/

Bovolenta, G., Politzer-Ahles, S., & Husband, E. (2016). Differential ERPs to local vs. global prediction failures. 29th CUNY Conference on Human Sentence Processing. Gainesville, United States.

Chen, J., & Chen, T. (in press). Mechanism and locus of tone 3 sandhi during Mandarin Chinese spoken word production. Journal of Chinese Linguistics.

Chen, Y., Shen, R., & Schiller, N. (2011). Representation of allophonic tone sandhi variants. Proceedings of Psycholinguistics Representation of Tone (satellite workshop to International Congress of Phonetic Sciences).

Chen, X. (2012). On the role of the underlying tonal representation in the phonological encoding stage of tone sandhi production of the Mandarin third tone. M.A. thesis, Guangdong University of Foreign Studies.

Chien, Y., Sereno, J., & Zhang, J. (2016). Priming the representation of Mandarin tone 3 sandhi words. Language, Cognition and Neuroscience, 31, 179 – 189.

Coltheart, M. (1980). Reading, phonological recoding, and deep dyslexia. In M. Coltheart, K. Patterson, & J. Marshall [Eds], Deep dyslexica (pp. 197-226). London: Routledge & Kegan Paul,.

Da, J. (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In Zhang, P., Xie, T., & Xu, J. [Eds.], Proceedings of the Fourth International Conference on New Technologies in Teaching and Learning Chinese (pp. 501-511). Beijing: Tsinghua University Press.

Fiorentino, R., Politzer-Ahles, S., Pak, N., Martínez-García, M., & Coughlin, C. (2015). Dissociating morphological and form priming with novel complex word primes: evidence from masked priming, overt priming, and event-related potentials. The Mental Lexicon, 10, 413-434.

Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. Psychological Review, 105, 251-279.

Gu, F., Li, J., Wang, X., Hou, Q., Huang, Y., & Chen, L. (2012). Memory traces for tonal language words revealed by auditory event-related potentials. Psychophysiology, 49, 1353-1360.

Hauk, O., Davis, M., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. NeuroImage, 30, 1383-1400.

Hunt, L., Politzer-Ahles, S., Gibson, L., Minai, U., & Fiorentino, R. (2013). Pragmatic inferences modulate N400 during sentence comprehension: Evidence from picture-

sentence verification. Neuroscience Letters, 534, 246-251.
Kouider, S., & Dupoux, E. (2009). Episodic accessibility and morphological processing: evidence from long-term auditory priming. Acta Psychologica, 130, 38-47.
Kutas, M., & Federmeir, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences, 4, 463-470.
Li, X., & Chen, Y. (2015). Representation and processing of lexical tone and tonal variants: evidence from the mismatch negativity. PLoS ONE, 10, e0143097.
McEnery, T., & Xiao, R. (2004). The Lancaster corpus of Mandarin Chinese. University of Oxford Text Archive.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. Clinical Neurophysiology, 118, 2544 – 2590.
Nixon, J., Chen, Y., & Schiller, N. (2015). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. Language, Cognition and Neuroscience, 30, 491-505.
Pisoni, D., & Levi, S. (2007). Representations and representational specificity in speech perception and spoken word recognition. In Gaskell, M. [Ed.], Oxford Handbook of Psycholinguistics (pp. 3-18). Oxford: Oxford University Press.
Politzer-Ahles., S. (manuscript). MMN experiments on lexicality of sandhi-derived allomorphs in Mandarin. Available at http://www.mypolyuweb.hk/~sjpolit/pubs/filedrawer/T3SlexicalityMMN.html
Politzer-Ahles, S. (2015). "Maybe" not all scalar implicatures are created equal. LSA Extended Abstracts.
Politzer-Ahles, S., & Connell, K. (submitted). Implicit priming of Mandarin third tone sandhi: a replication study. International Conference on Theoretical East Asian Psycholinguistics.
Politzer-Ahles, S., Fiorentino, R., Jiang, X., & Zhou, X. (2013). Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. Brain Research, 1490, 134-152.
Politzer-Ahles, S., & Gwilliams, L. (2015). Involvement of prefrontal cortex in scalar implicatures: evidence from magnetoencephalography. Language, Cognition and Neuroscience, 30, 853-866.
Politzer-Ahles, S., Schluter, K., Wu, K., & Almeida, D. (2016). Asymmetries in the perception of Mandarin tones: evidence from mismatch negativity. Journal of Experimental Psychology: Human Perception and Performance, 42, 1547-1570.
Politzer-Ahles, S., Xiang, M., & Almeida, D. (2015). "Before" and "after": investigating the relationship between temporal connectives and chronological ordering using event-related potentials. 7th Neurobiology of Language Conference. Chicago, United States.
Politzer-Ahles, S., & Zhang, J. (in press). Evidence for the role of tone sandhi in Mandarin speech production. Journal of Chinese Linguistics.
Pulvermüller, F. (1999). Words in the brain's language. Behavioural and Brain Sciences, 22, 253 – 336.
Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: the

mismatch negativity as a tool for studying higher cognitive processes. Progress in Neurobiology, 79, 49-71.

Schluter, K., Politzer-Ahles, S., & Almeida, D. (2016). No place for /h/: ERP investigation of English fricative place features. Language, Cognition and Neuroscience, 31, 728-740

Schröger, E., & Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. Neuroreport, 25, 3005 – 3008.

Shtyrov, Y. & Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. Neuroreport, 13, 521 – 525.

Smith, N., & Kutas, M. (2015). Regression-based estimation of ERP waveforms: I. The rERP framework. Psychophysiology, 52, 157-168.

Tucker, M., Politzer-Ahles, S., King, J., & Almeida, D. (2014). Agreement attraction in the neural language system. 20th Architectures and Mechanisms for Language Processing Conference (Edinburgh, United Kingdom) and 6th Neurobiology of Language Conference (Amsterdam, Netherlands).

Wiener, S., & Ito, K. (2015). Do syllable-specific tonal probabilities guide lexical access? Evidence from Mandarin, Shanghai and Cantonese speakers. Language, Cognition, and Neuroscience, 30, 1048-1060.

Wiener, S., & Turnbull, R. (2013). Mandarin word recognition: tone sandhi, accidental gaps and the lexicon. Poster presented at Linguistic Society of America.

Zhang, C., Xia, Q., & Peng, G. (2015). Mandarin third tone sandhi requires more effortful phonological encoding in speech production: evidence from an ERP study. Journal of Neurolinguistics, 33, 149 – 162.

Zhou, X., & Marslen-Wilson, W. (1997). The abstractness of phonological representation in the Chinese mental lexicon. In Chen, H. [Ed.], Cognitive Processing of Chinese and Related Asian Languages (pp. 3-27). Hong Kong: The Chinese University Press.

PROJECT FUNDING

3. Cost and justification

(a) Estimated cost and resource implications:
[Detailed justifications should be given in order to support the request for each item below]
(a maximum of 500 words for each box)

	Year 1 ($)	Year 2 ($)	Year 3 ($)	Year 4 ($)	Year 5 ($)	Total ($)

(A) One-line Vote Items

(i) Supporting Staff Costs

[please read Section 3(a)(A)(i) of the Explanatory Notes ECS2 carefully]

Types Monthly salary x Nos. x Months	Year 1 ($)	Year 2 ($)	Year 3 ($)	Year 4 ($)	Year 5 ($)	Total ($)
Research Assistant						$682,101
18,900 * 1 * 12	226,800					
20,790 * 1 * 12		249,480				
22,869 * 1 * 9			205,821			

Justification:

A Research Assistant/Associate will be recruited for this project. This project is labor-intensive, including an estimated 285 ERP sessions (4 experiments with N=55 each and one with approximately N=65), each of which may take around 2 hours. A Research Assistant/Associate who is well versed in running behavioral and ERP experiments with human subjects and in analyzing ERP data is indispensable for the timely completion of all experiments in this project. The duties of the Research Assistant/Associate include recruiting participants, running ERP experiments, analyzing ERP data under the supervision of the PI, and assisting as needed with preparing publications based on the data collected.

Year 3 only budgets for 9 months rather than 12 months, as the experiment planned for Year 3 is less labor-intensive (N=65 rather than around N=110).

(ii) Equipment Expenses
[please itemize and provide quotations for each item costing over $200,000]

Justification:

Quotation Provided: **Yes** ☐ **No** ☑

(iii) Outsourcing Expenses of Research Work Outside Hong Kong
[please itemize your cost estimation with justification and provide quotations for work costing over $200,000]

Justification:

Quotation Provided: **Yes** ☐ **No** ☑

(iv) General Expenses
[please itemize and provide quotations for services/purchase costing over $200,000]

FIRST YEAR: Participant fees: $33,000 EEG consumables -- Quik-Gel Electrolyte (128oz-bottles): $25,000. Open-access publishing fees: $17,500 SECOND YEAR: Participant fees: $33,000 EEG consumables -- Quik-Gel Electrolyte (128oz-bottles): $25,000. Open-access publishing fees: $17,500	75,500	75,500	45,500	0	0	196,500

THIRD YEAR:
Participant fees: $13,000
EEG consumables -- Quik-Gel Electrolyte (128oz-bottles): $15,000.
Open-access publishing fees: $17,500

Justification:

FIRST YEAR: 110 participants will participate in EEG experiments (the goal is N=48 for each of Experiment 1a and Experiment 1b, but I budget for several extra participants in case some participants need to be excluded because of data artifacts). Each participant will receive $250 compensation (about 3 hours) plus reimbursement for travel expenses up to $50 (Total: ($250+$50)*110 = $33,000).
EEG consumables: Quik-Gel Electrolyte (128oz-bottles). Total: $25,000.
Open-access publishing fees (estimated based on the average fee of discipline-relevant PLoS journals, Frontiers journals, Elsevier journals, and APA journals): $17,500

SECOND YEAR: 110 participants will participate in EEG experiments (the goal is N=48 for each of Experiment 1a and Experiment 1b, but I budget for several extra participants in case some participants need to be excluded because of data artifacts). Each participant will receive $250 compensation (about 3 hours) plus reimbursement for travel expenses up to $50 (Total: ($250+$50)*110 = $33,000).
EEG consumables: Quik-Gel Electrolyte (128oz-bottles). Total: $25,000.
Open-access publishing fees (estimated based on the average fee of discipline-relevant PLoS journals, Frontiers journals, Elsevier journals, and APA journals): $17,500

THIRD YEAR: 65 participants will participate in EEG experiments (the goal is N=48 for each of Experiment 1a and Experiment 1b, but I budget for several extra participants in case some participants need to be excluded because of data artifacts). Each participant will receive $150 compensation (about 2 hours) plus reimbursement for travel expenses up to $50 (Total: ($150+$50)*110 = $13,000).

EEG consumables: Quik-Gel Electrolyte (128oz-bottles). Total: $15,000.
Open-access publishing fees (estimated based on the average fee of discipline-relevant PLoS journals, Frontiers journals, Elsevier journals, and APA journals): $17,500

Quotation Provided: Yes ☐ No ☑

(v) Conference Expenses

One international conference per year to present results from this project.	20,000	20,000	20,000	0	0	60,000

Justification:

Standard expenses for conference travel to disseminate the output of this project at international conferences.

Sub-total for (A) (One-line Vote Items): $ 938,601

(B) Earmarked Items

(vi) Costs for Employment of Relief Teacher
[see Enclosure I for relief support under Humanities and Social Sciences Panel]

Rank

Monthly salary x Months

Justification:

Current Average Teaching Load: Total 0 classroom hours per academic year [please report UGC-funded programmes only]

(vii) High-performance Computing Services Expenses

Justification:

Quotation Provided: Yes ☐ No ☑

(viii) Research-related Software Licence /Dataset
[Please itemize and provide quotations for each item]

Justification:

Sub-total for (B) (Earmarked Items):	**$**	0
(ix) Total cost of the proposal (A) + (B)	**$**	938,601

(C) Deduction Items

Less :		
(x) Other research funds secured from other sources	**$**	0
Sub-total for C :	**$**	0
(xi) Amount requested in this application : (A) + (B) - (C)	**$**	938,601

(D) Academic Research related to Public Policy Developments

(xii) Percentage of the total cost of the proposal related to public policy developments ((A) + (B)) **[see Enclosure II for Support for Academic Research relating to Public Policy Developments]**	0%

(b) Declaration on the Equipment Procurement:

☑ **(i) No procurement of equipment is required**

OR

☐ **(ii) I declare that the equipment indicated in 3(a)(A)(ii) above is not available in the university**

OR

☐ **(iii) I declare that all or some of the equipment (please provide details in the following text box) indicated in Section 3(a)(A)(ii) above is available in the university but cannot be used by me in view of the following reasons (a maximum of 500 words)**

Reasons : (a maximum of 500 words)

(c) Declaration on employment of relief teacher:

☑ (i) No relief teacher is required

OR

☐ (ii) I declare that I currently do not hold any grant for employment of relief teacher of any on-going project under UGC/RGC funding schemes

OR

☐ (iii) I declare that I hold grant for employment of relief teacher of the following on-going project(s) under UGC/RGC funding schemes (excluding Humanities and Social Sciences Prestigious Fellowship Scheme (HSSPFS)) and undertake to submit the corresponding completion report(s) by 30 April 2017

(d) Declaration on high-performance computing services:

☑ (i) No procurement of high-performance computing services is required

OR

☐ (ii) I declare that the high-performance computing services indicated in Section 3(a)(B)(vii) above is not available in the university

OR

☐ (iii) I declare that all or some of the high-performance computing services (please provide details in the following text box) indicated in Section 3(a)(B)(vii) above is available in the university but cannot be used by me in view of the following reasons(a maximum of 500 words)

Reasons : (a maximum of 500 words)

(e) **Declaration on the research-related software licence / dataset:**

☑ **(i) No procurement of research-related software licence / dataset is required**

OR

☐ **(ii) I declare that the research-related software licence / dataset indicated in Section 3(a)(B)(viii) above is not available in the university**

OR

☐ **(iii) I declare that all or some of the research-related software licence / dataset (please provide details in the following text box) indicated in Section 3(a)(B)(viii) above is available in the university but cannot be used by me in view of the following reasons (a maximum of 500 words)**

Reasons : (a maximum of 500 words)

4. **Existing facilities and major equipment available for this research proposal:**
(a maximum of 400 words)

Department of Chinese and Bilingual Studies has a fully equipped EEG laboratory with a 128-channel Neuroscan Synamps amplifier, three 64-channel QuikCaps, an RF-shielded and sound-proofed booth, and software licenses for Neuroscan CURRY for data acquisition and STIM, E-Prime, and Presentation for experiment control. The department also has a departmental license for MATLAB (and the lab also has three standalone MATLAB licenses), which is what will be used for data preprocessing and statistical analysis.

5. Funds secured or to be secured

(a) Other research funds already secured for this research proposal:
[This amount will be deducted from the total cost of the project in Section 3 of Part II above.]

Source	Amount ($)

(b) Other research funds to be or are being sought for this research proposal.
[If funds under this item are secured, the amount to be awarded may be reduced]:

Source	Amount ($)

6. Particulars of PI

(Please check Proposal Update for latest CV)
(a) Investigator information:
Name and Academic Affiliation of Applicant:

	Name	Post	Unit/ Department/ University	Current Member of RGC Council as at the application deadline (Yes or No)	Current Member of RGC Subject Panel as at the application deadline (Yes or No)	Name of RGC Subject Panel
PI	Dr POLITZER-AHLES, Stephen	Assistant Professor	Department of Chinese and Bilingual Studies/The Hong Kong Polytechnic University	No	No	

(b) Curriculum vitae (CV) of Applicant.

[Please attach a maximum of two A4 pages in standard RGC format for attaching PDF documents or a maximum of 800 words for direct input in the text box) in the following format.]

i) Name:

ii) Academic qualifications:

iii) Previous academic positions held(with dates):

iv) Present academic position:

v) Previous relevant research work:

vi) Publication records [Please refer to ECS 2 Part II Section 6 for the format required by the RGC]:

vii) Others (please specify):

(c) Plan(s) for collaboration in this application:

[Indicate the role and the specific task(s) the PI and each Collaborator , if any, is responsible for.]

NA

(d) Number of hours per week to be spent by the PI in the proposal: 10 hour(s)

Stephen Politzer-Ahles

ACADEMIC QUALIFICATIONS

2013 Ph.D, Linguistics, University of Kansas
2011 M.A., Linguistics, University of Kansas
2008 B.A., Modern Languages and Literatures, Kenyon College

PREVIOUS ACADEMIC POSITIONS HELD

2015-2016 **Post-doctoral associate**, The University of Oxford. Faculty of Linguistics, Philology and Phonetics.
2013-2015 **Post-doctoral associate**, New York University Abu Dhabi. NYUAD Institute.

PRESENT POSITION

2016- **Assistant Professor**, The Hong Kong Polytechnic University. Department of Chinese and Bilingual Studies.

PREVIOUS RELEVANT RESEARCH WORK

Running behavioral experiments on tone sandhi in the United States and Beijing; running ERP experiments on tone sandhi and tone representation in Beijing and the United Arab Emirates

REPRESENTATIVE PUBLICATIONS WITHIN THE LAST FIVE YEARS

in press **Politzer-Ahles, Stephen**; & Zhang, Jie. Evidence for the role of tone sandhi in speech production. *Journal of Chinese Linguistics*. (Author 1 of 2)

2016 **Politzer-Ahles, Stephen**; Schluter, Kevin; Wu, Kefei; & Almeida, Diogo. Asymmetries in the perception of Mandarin tones: evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1547-1570. (Author 1 of 4)

2016 Schluter, Kevin; **Politzer-Ahles, Stephen**; & Almeida, Diogo. No place for /h/: ERP investigation of English fricative place features. *Language, Cognition, and Neuroscience, 31*, 728-740. (Author 2 of 3)

2016 Fiorentino, Robert; **Politzer-Ahles, Stephen**; Pak, Natalie; Martínez-García, María Teresa; & Coughlin, Caitlin. Probing the dynamics of complex word recognition: An ERP Investigation of the processing of novel compounds. *The Mental Lexicon, 10*, 413-434. (Author 2 of 5)

2013 **Politzer-Ahles, Stephen**; Fiorentino, Robert; Jiang, Xiaoming; & Zhou, Xiaolin. Distinct neural correlates for pragmatic and semantic meaning processing: An event-related potential investigation of scalar implicature processing using picture-sentence verification. *Brain Research, 1490*, 134-152. (Author 1 of 4)

ADDITIONAL PUBLICATIONS

2016 **Politzer-Ahles, Stephen**; Holliday, Jeffrey; Girolamo, Teresa; Spychalska, Maria; & Berkson, Kelly. Is linguistic injustice a myth? A response to Hyland (2016). *Journal of Second Language Writing, 34*, 3-8. (Author 1 of 5)

2015 **Politzer-Ahles, Stephen**; & Gwilliams, Laura. Involvement of prefrontal cortex in scalar implicatures: evidence from magnetoencephalography. *Language, Cognition, and Neuroscience, 30*, 853-866. (Author 1 of 2)

2013 **Politzer-Ahles, Stephen**, & Fiorentino, Robert. The realization of scalar inferences: context sensitivity without processing cost. *PLoS ONE, 8*, e63943. (Author 1 of 2)

2013 Lee, Hyunjung; **Politzer-Ahles, Stephen**; & Jongman, Allard. Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics, 41*, 117-132. (Author 2 of 3)

2013 Hunt, Lamar; **Politzer-Ahles, Stephen**; Gibson, Linzi; Minai, Utako; & Fiorentino, Robert. Pragmatic inferences modulate N400 during sentence comprehension: evidence from picture-sentence verification. *Neuroscience Letters, 534*, 246-251. (Author 2 of 5)

RESEARCH-RELATED AWARDS

2013 One-University Open Access Publishing Fund, University of Kansas
2013 CUNY Travel Award, CUNY Conference on Human Sentence Processing
2012 Doctoral Student Research Fund, University of Kansas
2012 IPA Student Award, International Phonetic Association
2011-2012 France Ingemann Scholarship, Linguistics Department – University of Kansas
2009 Linguistics Honors Fellowship, Linguistics Department – University of Kansas
2008-2009 Graduate Studies Scholarship, University of Kansas

CONSULTANCY

2016 Invited workshop on mixed-effects modeling in R [Linguistics, University College London], 18-20 May

GRANT REFEREEING

2013, 2015 *National Science Foundation (Linguistics panel)*
2013 *Swiss National Science Foundation*
2015 *FWO - Flanders*

JOURNAL REFEREEING

Bilingualism: Language and Cognition; *Brain and Language*; *Cognition*; *Developmental and Cognitive Neuroscience*; *Developmental Neuropsychology*; *Developmental Science*; *Frontiers in Psychology*; *Frontiers in Neuroscience*; *JASA Express Letters*; *Journal of Memory and Language*; *Language and Speech*; *Language, Cognition and Neuroscience*; *Neuropsychologia*; *Psychonomic Bulletin and Review*

DECLARATION OF SIMILAR OR RELATED PROPOSALS & GRANT RECORD

[Please refer to ECS2 for information required and implications for non-disclosure of similar or related proposals]

7. Re-submission of a proposal not supported previously

(a) Is this proposal a re-submission or largely similar to a proposal that has been submitted to but not supported by the UGC/RGC or other funding agencies?

Yes ☐ No ☑

If yes, please state the funding agency(ies) and the funding programme(s):

Reference No(s). [for UGC/RGC projects only]:

Project title(s) [if different from Section 1(a) of Part I above]:

Date(month/year) of application:

Outcome:

(b) If this application is the same as or similar to the one(s) submitted but not supported previously, what were the main concerns / suggestions of the reviewers then?

(c) Please give a brief response to the points mentioned in Section 7(b) above, highlighting the major changes that have been incorporated in this application.

8. Grant Record of Principal Investigator

Details of research projects undertaken and proposals submitted by the PI (in a PI/PC or Co-I/Co-PI capacity) including (i) completed research projects funded from all sources (irrespective whether from UGC/RGC) in the past five years; and (ii) on-going research projects funded from all sources (irrespective whether from UGC/RGC); (iii) terminated projects funded by UGC/RGC in the past five years; (iv) unsuccessful proposals or withdrawn projects submitted to UGC/RGC in the past five years; and (v) proposals pending funding approval. If you have any research project(s) / proposal(s) (not limited to the past five years) which is/are similar or related to this application, please include in the table below and provide an explanation on the differences between that/those

project(s)/proposal(s) and this application (a maximum of 400 words). [If you have difficulty in making the declaration, please explain.] Please add a new table for each project/proposal.

*** The PI is not required to report on the time spent in the capacity of Co-I in GRF / Joint Research Schemes projects.**

Major research output of previously funded projects

(For applications of Humanities and Social Sciences Panel only)

Major research output of previously funded projects (UGC/RGC and non-UGC/RGC sources) in descending chronological order, undertaken by the PI and each Co-I relevant to this application.
[Provide a summary (a maximum of 400 words in total) on the progress/ publications/ conferences/ student-training, etc. of the projects, with the relevant project reference no.]

Nil

ANCILLARY INFORMATION

9. Research Ethics/Safety Approval and Access to Government/ Official/ Private Data and Records

[Please refer to ECS2 Part II Section 9 for the responsibilities and implications]

(a) Research Ethics/Safety Approval

(i) I confirm that the research proposal ☑ **involves /** ☐ **does not involve human subjects.**

(ii) Please tick the appropriate boxes to confirm if approval for the respective ethics and/or safety issues is required and has been / is being obtained from the PI's university. PIs are encouraged to seek necessary approval (except for human research ethics (clinical)) before application deadline as far as possible

	Approval not required	Approval being sought	Approval obtained
(1) Animal research ethics	☑	☐	☐

	Approval not required	Approval being sought	Approval will be sought if funded
(2) Biological safety	☑	☐	☐
(3) Ionizing radiation safety	☑	☐	☐
(4) Non-ionizing radiation safety	☑	☐	☐
(5) Chemical safety	☑	☐	☐
(6) Human research ethics (non clinical)	☐	☑	☐
(7) Human research ethics (clinical)	☑	☐	☐

(iii) If approval is required by other authorities, please indicate *below* the names of the authorities and the prospects of obtaining such approval. If not applicable, please put down "N.A.".

NA

(b) Access to Government/ Official/ Private Data and Records

(i) Is access to Government / official / private data and records critical to the research proposal?

☐ **Yes**

☑ **No**

If approval is required, please indicate below the names of the agency(ies) of obtaining such approval.

(ii) Please tick in the appropriate boxes to confirm if approval for access to the related data/records has been / is being obtained from the relevant agency(ies). If approval has been obtained, please provide evidence.

List of agency(ies)	Approval not required	Approval being sought	Approval obtained

[Note: PIs are encouraged to seek necessary approval before application deadline as far as possible.]

10. Proposed reviewers

Note from the Secretariat:

At the December 2016 RGC meeting, it was agreed that the practice for applicants to nominate reviewers would be obsolete in view of the availability of RGC's sizable database of external reviewers and easily accessible resources on the internet. The reviewers nominated by the applicants in this exercise would not be adopted.

11. Release of completion report, data archive possibilities and public access of publications resulting from research funded by the RGC

(a) Is the proposed project likely to generate data set(s) of retention value?

Yes ☑ **No** ☐

If yes, please describe the nature, quantity and potential use of the data set(s) in future.

All EEG datasets generated in this experiment will be stored indefinitely (at MATLAB .mat files) on a project page hosted at the Open Science Foundation (http://osf.io) and/or GitHub. This will allow the PI and other researchers in the future to use these data for power analyses, follow-up analyses of the present research, meta-analysis, testing new statistical methods, and other things that further the goals of open data / open science.

(b) Are you willing to make the data set(s) available to others for reference twelve months after the publication of research results or the completion of this proposed project?

Yes ☑ **No** ☐

I understand that the RGC will release the completion report to the public and only considers data archiving requests after the completion of the RGC-funded project. The RGC has full discretion in funding the archiving requests. Data sets archived with RGC funds will require users to acknowledge the originator and the RGC. The originator will also be provided with copies of all publications derived from the use of the data.

I undertake to include in the project completion report the URL links to the university's repository or the publisher's websites so that the public could have quick and easy access to the manuscripts or journal articles. I will also consider to include in the research completion report the data repository where research data of the project could be accessed and shared, where appropriate.

I undertake that upon acceptance of a paper for publication,

(i) I will check whether the publisher already allows (A) full open access to the publisher's version, or (B) my depositing a copy of the paper (either the publisher's version or the final accepted manuscript after peer-review) in the university's repository for open access;

(ii) if both (i) (A) and (B) are not allowed, I will request the publisher to allow me to place either version in my university's repository for restricted access immediately upon publication or after an embargo period of up to twelve months if required by the publisher; and

(iii) subject to the publisher's agreement on (i) or (ii) above, I will deposit a copy of the publication in my university's repository as early as possible but no later than six months after publication or the embargo period, if any.

12. Education Plan, Dissemination Plan and Supporting documents

(A maximum of 20 words for each box to caption each uploaded pdf documents)

[Note: To ensure fairness in the independent assessment process, PIs should not disclose his/her identity or university in the Education Plan.]

Appendix 1: Education Plan (up to one A4 page)

Appendix 2: Dissemination Plan

EDUCATION PLAN

Research Assistant course fees: \$27,000

Research Assistant workshop fees: \$17,500

Undergraduate student helper training: \$20,000

Total: \$64,500

There are two potential educational components to this project: training the Research Assistant(s) and training possible undergraduate student helpers.

The Research Assistant(s) for this project will most likely be chosen from students who have recently completed a course in Psycholinguistics or Neurolinguistics. The ideal Research Assistant will be an undergraduate student who has expressed an interest in applying for graduate school—in my opinion, training a well-motivated student [who may go on to complete much more research in her future career] is just as important a part of the proposed project as the research itself. For a Research Assistant who has completed one of the abovementioned courses and not the other (e.g., if an exemplary student from Psycholinguistics agrees to work as a Research Assistant for this project, but has not taken Neurolinguistics yet), the PI will cover the course fees for the Research Assistant to take that course. As this is a three-year project, there could be as many as three different Research Assistants (although if the same person chooses to stay on as Research Assistant in this same project for two or three years, this will be preferred); thus, the total cost could be up to \$40,500:

\$4500 per credit × 3 credits per class × [up to] 3 Research Assistants = \$40,500

If the same student works as a Research Assistant for two or three years in a row, then the cost will be lower, as the same Research Assistant would not need to take the class again. In this case, the allocated funds could instead be used to send the Research Assistant to a workshop on data analysis, such as the ERP Boot Camp (http://erpinfo.org/the-erp-bootcamp); registration for this workshop is free, but housing costs would typically be about \$10,500 HKD (\$120 USD per night × 11 nights = \$1320 USD), and an air ticket from Hong Kong to Sacramento is estimated to be around \$7000. The most costly scenario under this scheme would be if the project had one student who served as an RA for one year (costing \$4500 × 3 = \$13,500 for course fees) and another who served as an RA for two years (costing \$13,500 course fees and approximately \$17,500 for workshop attendance), which would cost a total of \$44,500.

Undergraduate student helpers may also be trained to assist the Research Assistant with data collection. The rate for an undergraduate student helper is \$2500 per month. Data collection would not be ongoing constantly throughout the year; rather, I estimate that intensive data collection may last for one term (about 4 months) in a given year.

\$2500 per month × 4 months × 2 years = \$20,000

DISSEMINATION PLAN

The results of this project will be submitted to well-known international journals publishing in the areas of psycholinguistics and neurolinguistics, such as *Journal of Memory and Language*; *Language, Cognition, and Neuroscience*; *Journal of Experimental Psychology*; *Psychological Science*; *Cognition*; *Brain and Language*; *Journal of Cognitive Neuroscience*; *Laboratory Phonology*; *Quarterly Journal of Experimental Psychology*; *PLoS ONE*; or *Frontiers in Psychology*. Every effort will be made to publish the articles open access (and regardless of open access, preprints or author accepted manuscripts will always be made available through an institutional repository and/or on the PI's website, in accordance with the copyright agreement of the publisher), given that open access publishing has been shown to increase the readership and citation rate of articles, and thus can be expected to enhance the profile of the institution.

The results will also be presented at one or more of the major psycholinguistics- and neurolinguistics-themed international meetings, including *Society for the Neurobiology of Language*, *Architectures and Mechanisms in Language Processing*, *Cognitive Neuroscience Society*, *International Congress on Phonetic Sciences*, and *Tonal Aspects of Languages*.

PART III UNIVERSITY'S ENDORSEMENT AND DECLARATION OF RESEARCH ETHICS/SAFETY

(To be completed by the appropriate authority of the university. The university should confirm that it has evaluated and given support to the application before submission to the RGC.)

UNIVERSITY'S ENDORSEMENT

(* Please tick as appropriate in the boxes)

1. Staff Eligibility Requirement for ECS

I confirm that:

(a) the application has been evaluated and endorsed by the university for submission to the RGC;

(b) the PI meets fully the stipulated staff eligibility requirement for and is not debarred from applying for ECS grant;

[where the PI is newly appointed, the university has formally entered into a contract of service with him/her on or before the submission deadline of this funding exercise and the contract requires him/her to report for duty on or before 1 May 2017.]

(c) the PI is/will be employed in substantiation track / tenure track position;

(d) the applicant's statement at Part I Section 1 (a) (ii) is true;

(e) the applicant's declared teaching load at Part II Section 3 (a)(B)(vi) has been verified (only for the case where the applicant is seeking funding support for relief teacher);

(f) the applicant will have the number of hours per week as declared in Part II Section 6(d) to supervise the proposed project without prejudice to his / her existing commitment in other research work, teaching and administrative duties;

(g) the university will inform the RGC as soon as the PI ceases to be eligible to apply, receive or hold a ECS grant, and will withdraw the application / terminate the project if it is already funded and commenced;

(h) the university understands that the ECS grant, if given, will be withdrawn if the project does not start within one year of the announcement of funding result. The university should report to the RGC as soon as possible when a PI proceeds on no-pay leave/professional leave for a continuous or cumulative period exceeding 183 days within the project period;

(i) [for PI requesting Employment of Relief Teacher at Part II Section 3(B)(vi) only]

☐ the PI does not currently hold any grant for employment of relief teacher of any on-going project under UGC/RGC funding schemes

☐ the PI holds the grant for employment the relief teacher of the on-going project(s) at Part I Section 1(d) (excluding HSSPFS) and I shall ensure the PI to submit the corresponding completion report(s) by 30 April 2017

UNIVERSITY'S COMMITMENTS

2. Support to PI and Students

I confirmed that:

(a) ☑ the applicant will be permitted to be the primary supervisor of research student(s) (MPhil/PhD) during the project period.

(b) adequate supervision, research facilities and training provisions

☑ will

☐ will not

be in place to meet the need of RPg student(s) so employed under the research grant if this application is supported by the RGC.

☐ No RPg student will be trained in this proposed project

☐ Not applicable

(c) the research project under this ECS application

☑ is

☐ is not

in line with the role of the university.

(d) ☑ no outsourcing outside Hong Kong is required

☐ the PI's justification at Part II Section 3(a)(A)(iii) is reasonable and I support the PI's application for outsourcing the stated work stated outside Hong Kong

(e) ☑ no equipment is required

☐ the PI's declaration and reasons at Part II Section 3(b) are true and correct and I support the PI's request for this procurement of equipment

(f) ☑ no high-performance computing services is required

☐ the PI's declaration and reasons at Part II Section 3 (d) are true and correct and I support the PIs request for this subscription of high-performance computing resources

(g) ☑ no research-related software licence/dataset is required

☐ the PI's declaration and reasons at Part II Section 3(e) are true and correct and I support the PI's request for this procurement of research-related software licence/dataset

(h) ☑ this application does not include requests for purchasing normal academic equipment, computer, consumables, postage, fax, stationery, overseas telephone charges, and standard software licence / dataset known to be available or reasonably expected to be provided in the universities concerned.

(i) this application

☑ has

☐ has not (Please provide reasons for not scanning by anti-plagiarism software)

been scanned by anti-plagiarism software.

3. Research Ethics / Safety Approval and Access to Government/ Official/ Private Data and Records

(a) Research Ethics / Safety Approval

I have examined the research proposal and confirm that:

(i) ☑ **the research proposal involves human subjects**

☑ **and human research ethics (non clinical)**
- ☐ **approval has been obtained.**
- ☑ **approval is being sought.**
- ☐ **approval not required.**
- ☐ **exemption is being sought. / exemption has been obtained.**

☑ **and human research ethics (clinical)**
- ☑ **approval not required.**
- ☐ **approval is being sought.**
- ☐ **approval will be sought if funded.**

or

☐ **the research proposal does not involve human subjects.**

(ii) the approval of the appropriate authority(ies) is not required or has been/will be obtained in respect of the following :

	Approval not required	Approval being sought	Approval obtained
(i) Animal research ethics	☑	☐	☐
(ii) Biological safety	☑	☐	☐
(iii) Ionizing radiation safety	☑	☐	☐
(iv) Non-ionizing radiation safety	☑	☐	☐
(v) Chemical safety	☑	☐	☐

(b) Access to Government/ Official/ Private Data and Records

I have examined the research proposal and confirm that

(i) the approval of the appropriate authority(ies) is/are:

☐ **Required**

☑ **not required**

(ii) the approval of the appropriate agency(ies) has been/will be obtained in respect of the

following :

List of agency(ies)	Approval not required	Approval being sought	Approval obtained

For (a) and (b) above, except human research ethics (clinical), where such approval is required but has not yet been obtained, the university will ensure that it will be obtained without delay. The university understands that if no confirmation of such approval is provided to the RGC by 30 April 2017 , the RGC will regard this ECS application as being withdrawn and will stop further processing it.