Introduction

Heather Winskel

Traditionally, psycholinguistic research has predominantly focused on a small number of European languages. In more contemporary times, there has been a rapidly growing interest in investigating more diverse languages. By including a broader range of languages, we can build more comprehensive and representative universal models of psycholinguistic mechanisms and processes. The aim of the current volume is to contribute to this cross-linguistic endeavour and at the same time partially address this imbalance in research focus through reviewing some of the research conducted on the fascinating and varied languages of South and Southeast Asia. The people living in this large geographical region of South and Southeast Asia represent a large proportion of the world's population and speak a large number of diverse languages. South Asia comprises Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka and the Maldives. These areas are home to several hundred languages belonging to the Indo-European, Dravidian, Sino-Tibetan and Austro-Asiatic families. The area of Southeast Asia encompasses Vietnam, Cambodia, Thailand, Laos, Burma, Malaysia, Indonesia and southern China with at least five language families (Tai-Kadai, Sino-Tibetan, Hmong-Mien, Austroasiatic, Austronesian) spoken in this area (Enfield, 2005, 2011). The rich diversity and significance of these languages is reflected in the number of languages spoken; for example, in India it is estimated as being 438, Indonesia 719, and Vietnam 106 (Ethnologue). Moreover, languages such as Hindi, Malay or Indonesian are spoken by millions of people.

This volume on South and Southeast Asian psycholinguistics aims to highlight the rich diversity of languages and writing systems of this linguistically rich region. These languages are of growing interest to researchers as they differ significantly and in intriguing ways from the more commonly studied European languages in terms of grammar, lexicon, pragmatics or patterns in language usage, and written and spoken forms. Moreover, as a result of the historical genesis of psycholinguistics, most models and theories are based on a very limited number of languages and orthographies, predominantly European languages and Roman script; thus, it is essential to have data on a greater number of languages. As highlighted by Stoll and Lieven (Chapter 1),

'we have only very rudimentary information, sometimes only focusing on one linguistic feature, from approximately 1 per cent of the languages spoken today and within this 1 per cent there is a major bias towards acquisition studies of a very few languages'. In order to delineate between what mechanisms and processes are universal or common across languages and what are shaped by the characteristics of the particular language, it is essential to investigate the acquisition and processing of a more representative and diverse range of languages.

This volume was modelled on the very successful *Handbook of East Asian Psycholinguistics* series. However, in comparison to the languages covered by this series, namely Chinese, Japanese and Korean, much less research has been conducted and reached a wider international audience on the languages of South and Southeast Asia. The contributions in the current volume give a taste of the distinctive features of the languages of the South and Southeast Asian region. Even though this research represents only a tip of the psycholinguistic iceberg, it is important to acknowledge the significant contribution that the chapters in the current volume make to this enterprise. The volume covers topics and themes in first and second language acquisition, the development of reading and writing, diagnosis of language and literacy disorders in children and adults, and the relationship between language, brain, culture and cognition. It is hoped that this volume will form an invaluable resource to scholars and students interested in the languages of South and Southeast Asia as well as cognitive psychologists, linguists, educationalists, speech therapists and neuroscientists. An important goal of the current volume is to act as a catalyst to stimulate and encourage future research on these rich and diverse languages of South and Southeast Asia and their corresponding writing systems. First, we will provide a brief overview of some of the characteristics of the languages and writing systems of these two regions.

Characteristics of the languages and writing systems of South and Southeast Asia

The languages in the two geographic regions are very diverse in terms of both their linguistic characteristics and socio-cultural contexts. The Southeast Asian languages show striking convergence in terms of structure at all linguistic levels (Enfield, 2005, 2011; Matisoff, 1973, 2001). These languages are commonly analytic or isolating, thus, lack inflectional morphology (i.e. do not have agreement, case, gender/number/definiteness on noun phrases, tense-marking on verbs), have a tendency to be monosyllabic, have zero anaphora, rich inventories of sentence final particles, numeral classifiers and verb serialisation, and favour topic comment structure (Enfield, 2011). In addition, lexical tone is an important feature of many of these languages.

In marked contrast, Dravidian and Indo-Aryan languages have synthetic morphology, that is, several pieces of grammatical information are found within a word. Thus, these languages inflect for tense, gender and number, whereas most Southeast Asian languages lack this information and do not inflect at all. In place of inflectional morphology, the Southeast Asian languages typically utilise separate functor words or lexemes.

The rich scripts and writing systems of this region offer extremely fascinating opportunities for psycholinguistic research. Scripts of this region include Devanagari, Bengali, Urdu, Gujarati, Oriya, Kannada, Tamil, Telugu, Malayalam, Sinhala, Javanese, Thai, Lao, Khmer, Myanmar, Vietnamese and Tibetan. Many of the scripts of South and Southeast Asia share many features and characteristics as they are historically related and descendants of the ancient Brahmi script. They have been termed alphasyllabaries as they have hybrid characteristics of both alphabetic and syllabic scripts. Some of these scripts are used to write a diverse array of languages, for example Devanagari is used to write Hindi, Marathi and many other Indian languages. A common feature of many Indian scripts is *akshara*, which is basically an orthographic unit that represents sound at the level of the syllable, as well as marking constituent phonemes (see Joshi, Chapter 17; Bhuvaneshwari & Padakannaya, Chapter 18; Sircar & Nag, Chapter 19). Some of the Southeast Asian scripts do not have inter-word spaces (e.g. Thai, Lao, Khmer and Myanamar) and represent lexical tone orthographically (e.g. Thai, Lao, Vietnamese and Myanamar).

Scope and structure of the book

There are three sections to this volume. *Part I: Language acquisition*, which has two subsections – (i) Spoken language and (ii) Written language, *Part II: Language processing* and *Part III: Language and brain*. Some of the chapters are assigned somewhat arbitrarily to one of these sections, yet in reality, they cut across sections, disciplines and themes.

A number of recurring themes occur in this volume. One important recurring theme is bilingualism and multilingualism. Many children in this region are growing up in multilingual environments, learning to speak two or more languages and also learning to read and write in more than one language. Prime examples of multilingual and multicultural environments are Singapore and India. This is reflected in the chapters of the current volume. In multilingual Singapore, Susan Rickard Liow (Chapter 20) focuses on spelling development in bilingual children and Winston Goh and colleagues (Chapter 29) home in on the representation of languages in the memory of bilinguals and biscriptals. Obler and Paplikar (Chapter 31) review case and group studies conducted on multilingual individuals with aphasia in South

and Southeast Asia. In India, Nair, Ravi, Bhat and Chengappa (Chapter 11) discuss the fast mapping of novel words in bilingual and multilingual children and Bhat and Chengappa (Chapter 35) review research conducted on bilingual aphasia in this multilingual milieu.

An additional recurring theme is the importance of developing diagnostic language assessment instruments in the local language(s) spoken in the South and Southeast Asian region (Razak, Chapter 12; Jin, Razak, Wright & Song, Chapter 13; Nair *et al*., Chapter 11; Joshi, Chapter 17). The early identification of children with language, reading and communication problems, using appropriately designed assessment instruments in the child's native language, is crucial so that timely intervention programmes can be implemented.

Chinese has had a huge impact in the Southeast Asian region. It is a widely spoken and significant language in Southeast Asia. The important role that Chinese plays in the Southeast Asian region is reflected in the chapters by Jin *et al*. (Chapter 13) in Malaysia and Rickard Liow (Chapter 20), Goh, *et al*. (Chapter 29) and Oh (Chapter 34) in Singapore as well as the chapter on discourse pragmatics in Mandarin Chinese by Zhou (Chapter 3). Tone languages represent a large proportion of the spoken languages of the world (Yip, 2002), and yet lexical tone is an understudied feature in psycholinguistic research. In the current volume Kitamura (Chapter 2), Abramson (Chapter 21), Xu Rattanasone, Attina, Kasisopa and Burnham (Chapter 22) and Gandour and Krishnan (Chapter 32) all review research conducted on lexical tone and tonal languages.

Psycholinguistic methods

Psycholinguistic research, as represented by the current volume, includes a range of methodologies and techniques, naturalistic observations and case studies, as well as more experimental methodologies, which all play an invaluable role in increasing our understanding of psycholinguistic mechanisms and processes of these languages. There have been rapid advancements in the development of the new technologies available to psycholinguists. This volume includes chapters that highlight the use of brain imaging and eye-tracking technologies to gain greater insights and understanding of the psycholinguistic and neural correlates of behaviour.

Part I: Language acquisition

(i) Spoken language

Even though the languages of the world pose a diverse set of challenges for young children in terms of phonological, morphological and syntactic structure, typically developing children throughout the world rapidly become

competent and effective communicators in their particular language. There are striking commonalities in language development across diverse languages, but also intriguing language-specific effects on language acquisition.

In the introductory chapter to this section, Sabine Stoll and Elena Lieven (Chapter 1) discuss the major theoretical approaches and methods used in language acquisition research. They also examine the role of variation in terms of language structure, contexts of child-rearing, and individual differences. Subsequently, they focus on three prominent areas of research: early vocabulary acquisition, spatial cognition and acquisition of argument structure, which are used to address directly the question of whether children are geared by innate predispositions or whether they adapt to language-specific structures and categories from an early age.

As stated by Stoll and Lieven, 'one of the major challenges is to disentangle linguistic and cultural variables'. One way this can be achieved is by studying typologically distinct and diverse languages. The complex interplay between language-specific features, culture and input in language acquisition is illustrated by the chapters by Kitamura (Chapter 2), Zhou (Chapter 3), Ratitamkul (Chapter 4) and Shirai (Chapter 5). Kitamura and Zhou both discuss the important dual role of caregiver–child interactions in language acquisition and socialisation of the child. Kitamura discusses the role of infant-directed (ID) speech in tonal and non-tonal languages. A unique blend of features characterises ID speech in a diverse number of languages. However, there are also cross-cultural variations that reflect culture-specific beliefs and values. For example, Australian English and Thai mothers adapt their speech differently depending on gender of the child. Zhou (Chapter 3) focuses on pragmatics or patterns of language use between Mandarin-speaking mothers and their children. This chapter highlights the influence that culture has on interaction patterns between caregivers and their children. There are shared commonalities in interaction patterns across cultures but also culture-specific patterns shaped by the cultural beliefs and goals of the caregivers.

Telling stories or telling others about things that have happened to us is an integral part of everyday conversation in all cultures. In order to create a coherent narrative, a speaker needs to sequence events in a comprehensible order and refer appropriately to characters occurring in the narrative, while concurrently observing the discourse norms and conventions of the particular culture. Ratitamkul (Chapter 4) examines the use of referential forms by Thai children when telling the famous *Frog Story* (Mayer, 1967). She found that Thai children's use of referential forms to create coherence in discourse appears relatively late (around 9 years old), conforming to research findings on other languages.

Shirai (Chapter 5) reviews research on the acquisition of tense-aspect markers in some Asian languages: Japanese, Korean, Mandarin Chinese,

Vietnamese and Thai. The pattern of tense-aspect acquisition has been of great interest to researchers in relation to the issue of nature versus nurture (Shirai, 2009). As an alternative to Bickerton's innate bioprogram, Shirai and colleagues have proposed that the pattern of acquisition observed can be accounted for by the distribution of tense-aspect morphology occurring in the input to the child in conjunction with the cognitive processes involved in children's early category formation. Based on a review of the data from research conducted on these Asian languages, he concludes that input frequency gives a more plausible account than innate predisposition (i.e. a bioprogram) for the patterns in acquisition observed. Individual and task differences are also important additional factors to consider. In order to gain greater insights into the nature of universals and particulars of language acquisition, future research needs to further investigate the acquisition of tense-aspect systems in Asian languages using multiple tasks and methods.

Numeral classifier systems are a common feature of Southeast Asian languages. They are typically used in counting or referring to objects. In numeral classifier languages, objects are categorised into classifier categories based on their physical, conceptual and functional properties. The categories used by different numeral classifier languages have both shared and idiosyncratic characteristics. The complex semantic nature of numeral classifier systems poses significant challenges to young language learners. In the current volume, the numeral classifier systems of Malay (Salehuddin, Chapter 6) and Vietnamese (Tran, Chapter 7) form an interesting comparison. They categorise everyday objects and items in both common and quite distinct ways. The acquisition of Malay numeral classifiers is a relatively delayed and prolonged process in comparison to Vietnamese and other classifier languages. This appears to be due to the lesser degree of obligatoriness of this grammatical category in Malay.

Aman (Chapter 8) provides an account of the acquisition of wh-questions of children who speak colloquial Singapore Malay. Malay, in contrast to European languages such as English, German, French or Spanish, allows questions with both in-situ and moved wh-words. European languages lack the latter in-situ question form. Thus, this study presents an alternative perspective on the acquisition of wh-question forms.

Marathi, an Indo-Aryan language, is the official language of the state of Maharashtra and is spoken by over 70 million people. Moreover, Marathi is a richly inflected language. Pitale and Sarma (Chapter 9) investigate the acquisition of inflectional morphology of plural forms in Marathi using Berko's famous *wug* task. In Marathi, there is a complex interplay between grammatical gender (whether a noun is masculine, feminine or neuter) and the final segment of the noun in determining the appropriate plural form to be

used. This presents interesting challenges to the young language learner. Results are found to parallel those on the acquisition of English and German and indicate that morphological rules are induced by children. Morphological acquisition is governed by both universal principles of acquisition (e.g. developmental sequence, regularity of affixation and conditioning) and by specific features of the actual target language (e.g. number of nominal classes or genders, and phonotactics).

Tamil is a Dravidian language, primarily spoken in the Indian subcontinent as well as in Sri Lanka, Singapore and Malaysia and by emigrant communities in many other countries. Similar to Marathi, Tamil is a morphologically rich language. The functions that are expressed by the verbal complex are structurally and functionally highly complex and consequently pose significant challenges to language learners. Sarma (Chapter 10) traces the acquisition of this complex verbal morphological system by a young child. Verbs are the key determinants of argument structure, and hence, particularly important to investigate in the field of language acquisition.

Nair *et al*. (Chapter 11) discuss the fast mapping of novel words in bilingual and multilingual children. Fast mapping is a term coined to explain how children acquire vocabulary so rapidly during their early developmental years. The traditional perspective is that fast mapping is a process that enables children to create lexical representations for the unfamiliar words they encounter. However, very few studies have investigated fast mapping skills in children growing up in bilingual and multilingual contexts. This research can increase our understanding of children's word-learning skills in each of the languages spoken and how lexical representations are created and stored in the bilingual child's mental lexicon.

Razak (Chapter 12) reviews studies conducted on the acquisition of Malay language. There is a great need for further research investigating the language development of Malay-speaking children. This research can be used to determine typical developmental patterns and milestones in language development in Malay. These milestones of language development can then act as indicators to identify developmentally delayed and at-risk children, so that early intervention programs can be implemented.

Following on from this theme, Jin *et al*. (Chapter 13) discuss the issues involved in developing language assessment tools in two of the languages spoken in multilingual Malaysia, Malay and Chinese. Ascertaining the baseline norms and age-related language capabilities of normally developing Malay- and Mandarin-speaking children is the first step in developing language assessment tools. The overall aim is to provide a comprehensive, normative account of the acquisition of Malay and Chinese, so that children with communication problems can be identified at an early stage.

(ii) Written language

As an introductory chapter to this section, Brian Byrne and colleagues (Chapter 14) give a comprehensive review of research on learning to read in European languages. English, in particular, is often used as a benchmark for literacy research even though it can be considered to be an 'outlier' orthography due to its high degree of inconsistency or irregularity (Share, 2008). As European languages have historically played a seminal role in literacy research, there is an expectation that research on other languages will make comparisons with these well-studied languages and orthographies. As Byrne, Samuelsson and Olson point out, our perspective on reading and writing development could have been quite different if the focus of research had evolved around Chinese or Hindi, for example.

The chapters that follow review reading and writing development in a range of languages: Thai (Winskel, Chapter 15), Malay/Indonesian (Winskel & Lee, Chapter 16), Kannada (Joshi, Chapter 17), Tamil (Bhuvaneshwari & Padakannaya, Chapter 18), Bengali (Sircar & Nag, Chapter 19), and spelling development in bilingual children in Singapore (Rickard Liow, Chapter 20). This region offers rich opportunities for research on literacy acquisition in different scripts and children learning to read in more than one language and script.

Thai, a tonal language, has its own distinctive alphabetic script that shares some common characteristics with Indic writing systems, due to common origins. The particular challenges this distinctive orthography poses to beginning readers and writers of Thai is discussed by Winskel (Chapter 15). Thai also has syllabic characteristics as it has implicit vowels for some consonants. Consonants are written in a linear order, but vowels can be written above, below or to either side of the consonant as full letters or diacritics, and commonly combine across the syllable to produce a single vowel or dipthong. There is a high level of consistency of mapping between phonemes and graphemes but there are multigrapheme to phoneme correspondences; consequently, spelling development lags behind reading in Thai. Thai is predominantly a monosyllabic language with words with very similar spellings, often varying in just one letter or the tone of the syllable, which adds to the challenges of learning this particular orthography.

A variety of the Malay language is spoken in four Southeast Asian countries, namely Indonesia, Malaysia, Singapore and Brunei (Prentice, 1987). Winskel and Lee (Chapter 16) review research on learning to read and write in Malaysian/Indonesian. Malaysian/Indonesian language provides an interesting case study as it uses the same Roman script as English, but in contrast has a high degree of orthographic transparency. Furthermore, the syllable is a highly salient unit as it is predominantly bi- and multisyllabic and has a simple syllable structure with clear syllable boundaries.

Joshi (Chapter 17) describes the characteristics of Kannada language and its orthography. Current research in reading development and reading breakdown is reviewed. Kannada is a polysyllabic agglutinative language with numerous inflections. The script of Kannada is an alphasyllabary and thus lends an interesting contrast between alphabetic languages like English and syllabic languages like Japanese Kana. Based on current evidence, Kannada seems to resemble English more than Japanese Kana. Kannada orthography has a fairly consistent relationship between the written unit, *akshara*, and the sound. Even though the orthography is transparent, literacy acquisition is still a difficult task for various reasons. The complex alphasyllabic nature of the akshara of Kannada is a contributing factor. In addition, poor teaching methods, large class sizes and lack of books in school and at home exacerbate the problem. It is important that standardised assessment materials exploiting the nature of Kannada aksharas are developed.

Bhuvaneshwari and Padakannaya (Chapter 18) discuss the distinctive characteristics of Tamil language and its orthography and the challenges it poses to young readers as well as the interesting comparisons that it makes with other orthographies. Tamil script is a derivative of Brahmi script and its southern version, Grantha script. All elements of Tamil akshara are written as in line characters in contrast to other *akshara*-based scripts where characters can occur above or below the main text line. Tamil also has a relatively opaque or deep orthography in contrast to other Indic scripts, as it has one-to-many mapping of grapheme to phonemes. In addition, Tamil is diglossic with a marked difference between colloquial spoken language and more formal spoken and written forms. These distinctive characteristics pose significant challenges to the novice reader. A proficient reader needs to refer to sentence context and prior lexical knowledge to effectively read Tamil.

Sircar and Nag (Chapter 19) discuss the nature of mapping between *akshara* and the phonological syllables in Bengali. They examine the relative difficulty in word reading where *akshara*–syllable mismatches occur, and the ways in which akshara–syllable mappings impact phonological processing. Children's responses on word reading, syllable and phoneme processing tasks indicate that different types of *akshara*–syllable mappings are read and processed in different ways. These findings are one example of how specific properties of the spoken language and the writing system interact and influence the process of learning to read.

In the final chapter of this section, Rickard Liow (Chapter 20) investigates the processes that contribute to spelling development in English for three groups of bilingual children (English–Mandarin, Mandarin–English and Malay–English) in Singapore. A series of empirical studies explore the processes that contribute to spelling in English for these three subtypes of bilingual children who all follow the same curriculum. Early spellings are

speech-based (Jalil & Rickard Liow, 2008), and the three main languages (English, Mandarin and Malay) are dissimilar in terms of phonology as well as orthography. These differences in the characteristics of the languages spoken and scripts learned by the different subtypes of bilinguals, influence both phonological representation and spelling development.

Part II: Language processing

Lexical tone is a common characteristic of the languages of Southeast Asia and yet it is psycholinguistically an understudied feature. The languages of Southeast Asia can be categorised into tonal languages or voice-register languages as well as languages that fall into neither category. Arthur Abramson (Chapter 21) presents an overview of acoustic and physiological research on tones and voice registers in the languages of Southeast Asia with particular attention to studies of both production and perception. These languages include not only official national languages with their regional and social dialects but also the languages of numerous ethnic minorities. Examples of tonal languages are Thai, Lao, Vietnamese and Burmese and examples of voice-register languages are Chong, Suai, Kh'mu and Mon. Languages such as Malay, Indonesian and Cambodian, at least in their standard dialects, belong to neither category. Xu Rattanasone and colleagues (Chapter 22) continue this theme by examining a number of methods by which tones can be described and compared, ahead of three examples of research using variations of a relatively recent method, that of tone space mapping.

Mishra (Chapter 23) investigates sentence generation during scene-viewing in Hindi using eye-tracking technology. Eye movements offer real-time data related to cognitive processing and are particularly useful in studying language processing. Mishra presents and discusses the results from a recent study of sentence generation in children and adults in Hindi. The findings are discussed in relation to theories of language production and the multisensory interaction between language and vision.

Kim and Davis (Chapter 24) review some of their recent experimental findings with both developing and skilled Thai readers and make comparisons with results from English readers. They review findings from tests that range from a short-term priming technique that probes print-to-sound mapping to examining the metalinguistic skills associated with the awareness of lexical tone. Patterns of results were found that are consistent with those for English (problems in processing print-to-sound connections appear to be a common feature of reading difficulties) but also results emerged that are most likely due to the specifics of Thai.

Tsai (Chapter 25) and Reilly (Chapter 26) both use eye-tracking technology to investigate eye movements when reading Thai and Chinese. Both Thai and