Chapter 7: The Cost of Fluency

Kim Kirsner

Granted that fluency in the dominant language of a country is a *necessary* if not *sufficient* condition for integration, questions about the cost of fluency are pertinent. The aim of this chapter is to determine the amount of engagement required to achieve fluency in a second language.

Measurement of second language skills depends critically on the test procedure, and the target problem. Many procedures involve de-contextualised tasks, and they therefore fail to capture the speaker's fluency under natural language conditions. Other procedures rely on subjective judgement, and they are correspondingly unreliable. The procedure outlined here involves measurement of fluency for short natural language samples, and provides a range of estimates of the cost in time of first language fluency in a second language.Engagement was assessed via a structured and in-depth interview designed to provide an estimate in hours of each participant's exposure to his or her second language. The fluency procedure was designed to provide objective and semi-automatic measurement of pause duration, and effective information transmission was also calculated. We have recently implemented and tested a fully automatic procedure.

We describe a pilot study involving 24 English-Italian or Italian-English bilinguals in Australia. The final study will involve 400 people including one EFL and eight ESL groups. The interview yielded second language practice estimates from 250 to 100,000 hours for all forms of engagement in second language activities. We obtained five-minute speech samples from each participant in English and Italian. Our results indicated between 8,000 and 80,000 hours of engagement are required before second language skills meet first language levels.

The discussion focused on questions about technical developments required to transform the language learning equation.

Fluency and integration are complex concepts. Fluency in the language of the host country is certainly critical to effective integration, however other factors including family upbringing, access to local customs and traditions, emotional attachments, values and religion will also enter the equation. Fluency is also complex, particularly where spontaneous speaking is concerned. It reflects motor, cognitive and vocabulary capabilities where these variables will in turn reflect the amount of engagement the individual has had with the host culture. The core concepts of integration and fluency are, we assume, associated to some

extent, however the extent of that association and the causal links between the concepts are likely to be variable and unpredictable. The landscape of the relationship between integration and fluency is further compromised by the role of identity.

The aim of this project is to take a first and tentative step down the road to empirical measurement of a sample of these variables. In the pilot study described below we have concentrated on the relationship between Engagement and Fluency. Using a new and objective approach to the measurement of fluency – to what extent is fluency moderated by engagement in the host language?

The extent to which immigrants adapt to and integrate in their host society depends on a variety of individual, social, educational, emotional and cultural variables. One of the most potent of these variables involves the extent to which an individual has mastered the language of his or her host community. Mastery is a complex issue however, and involves many dimensions of communication. Here we focus on what is arguably the most demanding task in second language learning, language production.

Language production reflects both *competence*, an individual's knowledge of his or her own language, and *performance*, the transformation of competence into everyday speech[1]. In our approach we have focused on performance, and, even more narrowly, fluency. In this study consideration is restricted to just two dimensions of fluency, the amount of time required to deliver one *Correct Information Unit*, a measure that excludes material that is unintelligible, ungrammatical or not appropriate to the communication task, and *pause duration*, periods of silence during continuous speech.

Design

The results described here involve speech samples from eighteen English-Italian and six Italian-English speakers[2]. The speakers ranged in age from 17 to 56.

The participants were selected in order to provide a wide cross-section based on the extent of the participant's engagement in English. In practice, the range for the participants described in this chapter is from 250 hours for a recent arrival in Australia to 100,000 hours for one participant who left Italy 40 years ago.

1 Chomsky, N (1965). *Aspects of the Theory of Syntax*, MIT Press, Cambridge (Mass.).

2 Bujalka, H (2006). A Quantitative Assessment of the Development of Fluency in English/Italian and Italian/English. Unpublished Thesis, University of Western Australia, School of Psychology.

Each participant completed a questionnaire in collaboration with an interviewer in order to determine; first, the extent of his or her engagement in language activities involving all of her languages, and, second, his experience of, attitudes to, and integration with the Australian community.

Each participant subsequently provided five speech samples, one in his or her first language, and the balance in English. The samples were provided in response to questions about everyday topics. The following is a summary:

Question 1: First Language Response: Imagine your ideal holiday in the country of your first language (eg, Arabic, Italian, Greek, etc). Imagine a place you have never actually visited. Where would you go, what would you do there; and what would you expect to see?

Question 2: English Second Language Response: Imagine your ideal holiday in Australia. Imagine a place you have never actually visited. Where would you go, what would you do there; and what would you expect to see?

Question 3: English Second Language Response: Think about what you have done over the last two or three weeks. Where have you been? What did you see or hear that was new? What was difficult and what was easy? What did you enjoy and what did you find boring?

Question 4: English Second Language Response: Think about a typical day in your life. Think about a typical social event, a picnic, an outing, a visit to the beach, a party, or a meal for example, a visit to your favourite club or pub.

Question 5: English Second Language Response: I would be most grateful if you could tell me about your experience of Australia? How do you feel about Australia? Have you been accepted? Have you experienced difficulties in regard to education or work that appear to stem from the fact that English is not your first language? Just give me a general impression of your experience please.

Results

Estimation of second language engagement

A structured interview was designed and implemented to estimate the amount of practice each participant had had in English or Italian. The interviewer followed a chronological structure, and included questions about language use during each phase of the interviewee's life. The interviewees were specifically asked to estimate the amount of time associated with each language for each

phase. The structure of the interview included provision for questions about the interviewee's home, school, family, social, working and holiday roles as well as language instruction. The interview data was used to estimate engagement (in hours) for each participant.

Transmission of information

The amount of information conveyed in a message is not tied directly to the temporal dynamics of language production[3]. In Broca's aphasia for example, speech is often slow and difficult to follow but analysis can nevertheless reveal that a considerable amount of information has been transmitted. In Wernicke's aphasia on the other hand, fluency including the number of words spoken per minute may be very high, but the amount of information conveyed very low, as the actual words sound like 'gibberish'.

We therefore adopted an independent measure of communicative efficiency. The measure, based on work by Nicholas and Brookshire[4], involves the concept of a Correct Information Unit. The Correct Information Unit count for a speech sample is restricted to *words* that are intelligible, and accurate, relevant and informative relative to the eliciting stimulus. For comparative purposes we divided speaking time by Correct Information Units to give *seconds per Correct Information Unit*.

Figure 1 shows the impact of practice on Communicative Efficiency. The dashed function is a measure of Communicative Efficiency for the speaker's first language. On average, each unit requires about 0.4 seconds, and first language performance is more or less indifferent to practice in the speaker's second language. However, while speakers produce about 2.5 units per second in their first language, second language performance commences at about one unit every two seconds following a few hundred hours of second language practice and it does not intersect with first language performance until somewhere between 10,000 and 100,000 hours of practice. Communicative Efficiency in second language production evidently involves a significant outlay of time even for cognate languages such as English and Italian.

3 For example, Ano, K (2002). Relationship between fluency and accuracy in spoken English of High School Learners, Step Bulletin 14: 39-49.

4 Nicholas, L E and Brookshire, R H (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research*, 36: 338-350.

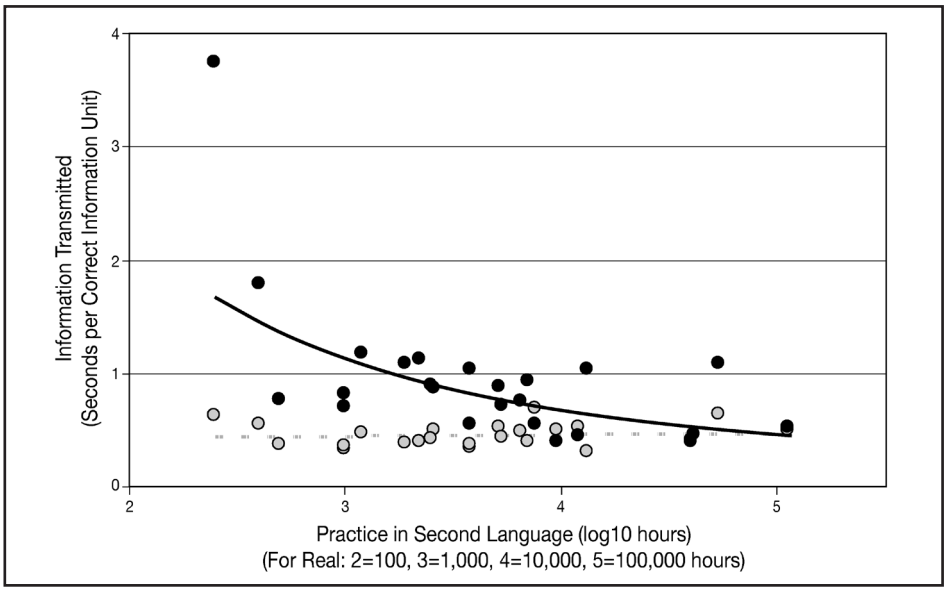

Figure 1 Relationship between practice and information transmitted

Role of fluency in second language learning

Measurement of the temporal dynamics is more challenging. Consideration of work by Lennon and Clark and Clifford[5] indicated that fluency can be used to refer to frequency of pauses, the duration of pauses, utterance length, semantic density, situational appropriateness and several other variables. There is some consensus however, that it refers to *performance* rather than *competence*[6]; that it is a measure of language as it is processed in real time[7]; and that it reflects interactions among a number of processes[8]. Uncertainty about the measurement of fluency extends to the measurement of second language skills. The assessment procedure for immigrants to Australia assigns individuals to bands. To achieve a particular band an individual must demonstrate mastery of specific features associated with a particular band. Is it possible to develop an objective procedure for measuring fluency?

One of our objectives is to develop a measurement system that is not only objective but automatic as well. We have developed acoustic and mathematical procedures[9], and we are currently refining an automatic implementation. We have

5 Lennon, P (1990). Investigating Fluency in EFL: A quantitative approach, *Language Learning*, 40: 387-417; Clark, JLD and Clifford, R (1988). 'The PSI/ILR/ACTFL proficiency scales and testing techniques: development, current status and needed research'. *Studies in Second Language Acquisition*, 10: 129-147.

6 Chomsky, N (1965). *Op cit*.

7 Schmidt, R (1992). Psychological mechanisms underlying second language fluency. *Studies inSecond LanguageAcquisitio*n, 14: 357-385.

8 Lennon, P (1990). *Op cit*.

9 For example, Kirsner, K, Dunn, J, Hird, K, Parkin, T and Clark, C (2002). 'Time for a Pause...', 9th International Conference on Speech Science, Melbourne; Kirsner, K, Dunn, J and Hird, K (2005). Language Production: a complex dynamic system with a chronometric footprint. *Proceedings of the 7th International*

however applied an interim stage of the procedure to a variety of communication problems including simultaneous interpretation and the acquisition of second language fluency

The temporal dynamics of language production are very complex, similar in some respects to output from an Electrocardiogram (EKG). Spontaneous speaking is composed of alternating periods of silence and speech. But there are several vexing measurement problems. The first of these has been recognised since the pioneering work of Lounsbury, and Goldman-Eisler[10]. The problem is that silence emanates from three or possibly more qualitatively different sources, including 'breathing', 'articulation' and 'cognition'.

The second problem is that pause duration distributions are massively skewed, precluding the use of arithmetic means to assess pause duration in natural language. This problem was first recognised by Quinting[11] but it has rarely been recognised in subsequent research.

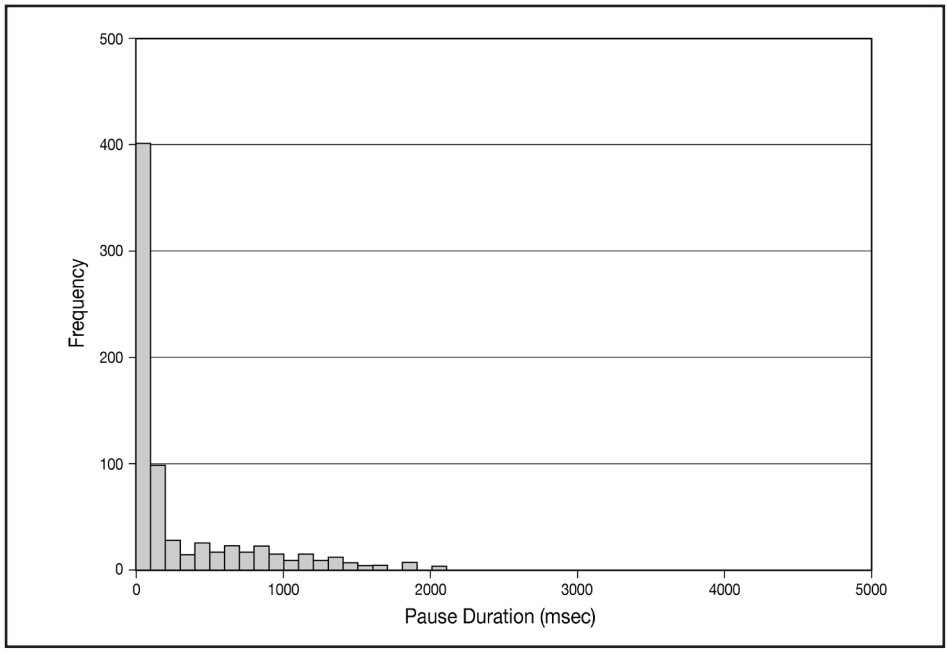

Figure 2 Typical pause duration distribution in real time

Conference on Cognitive Systems, New Delhi (India); Kirsner, K, Dunn, J J, Hird, K and Hennesy, N (2003). Temporal co-ordination; the lynch-pin of language production. In Palethorpe, S and Tabian, M (eds.), *Proceedings of the 6th International Seminar on SpeechProductio*n, Macquarie University, CD-ROM: 19-24.

10 Lounsbury, FG (1954). Transitional probability, linguistic structures. In C Osgood and T A Sebeock (eds.), *Psycholinguistics: a survey of theory and research problems*: 93-101, Waverley Press, Baltimore (MD); Goldman-Eisler, F (1968). *Psycholinguistics: Experiments in spontaneous speech*, Academic Press, New York. Goldman-Eisler, F (1968). *Psycholinguistic*s, Academic Press, London/New York.

11 Quinting, G (1971). *Hesitation phenomena in adult aphasic and normal speech*, Mouton, The Hague.

Figure 2 shows the pause duration distribution for one three-minute speech sample. The figure indicates frequency of occurrence for each bin. Hand analysis of pause duration is a challenging task of course, hence our motivation to produce an automatic system. Preparation of the figure involved reference to two hundred and thirty-nine 25 msec 'bins' between 25 msec and 6,000 msec or six seconds. As depicted in Figure 2, the vast majority of the hand-measured pauses fall into the first (26-50 msec), second (51-75 msec) or third bin (76-100 msec).

The assumption that two distributions are involved is not obvious in Figure 2. In Figure 3 however, following natural log transformation, the presence of two distributions is apparent[12].

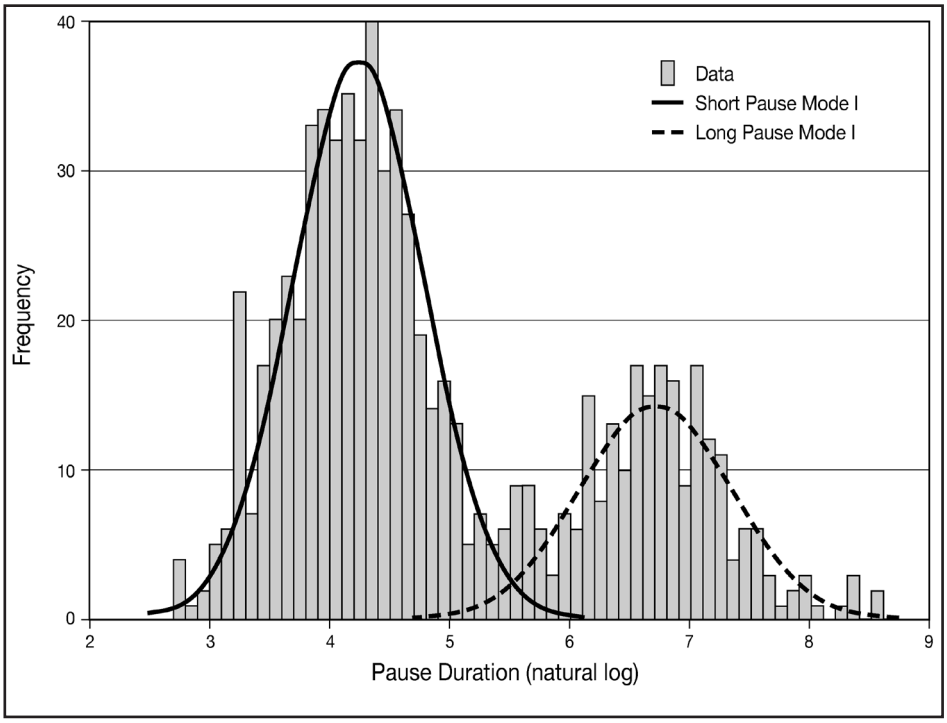

Figure 3 Typical pause duration distribution in natural log

The modes of the two-pause duration distributions fall at about log 4.1 (60 msec) and log 6.2 (600 msec) respectively. Recent work by our group indicates that most, although perhaps not all, the pauses in the short pause distribution constitute what Goldman-Eisler[13] would have called 'articulation' pauses, and that very few of the pauses in the long pause distribution fit the notion of

12 Kirsner, K *et al* (2002). *Op cit*.
13 Goldman-Eisler, F (1968). *Op cit*.

an 'articulation' pause. As a working hypothesis, it is our assumption that the short and long pause distributions reflect articulation and extra-articulation sources of variance respectively.

The solution to the measurement problem is apparent in Figure 3. The two distributions are approximately lognormal, and it is clear that they intersect at about log 5.5 or 250 msec, the value adopted by Goldman-Eisler to exclude articulation pauses from her analyses.

One more problem merits consideration. Complex biological systems are variable and unstable. Everything from nerve transmission time to conceptual planning varies from person to person, and occasion to occasion, and the actual point of intersection between the two distributions can fall anywhere between 100 and 400 msec for different individuals. Thus, although Goldman-Eisler[14] chose an appropriate mean value for the threshold – to exclude articulation pauses – adoption of the same value for each and every speech sample or participant inevitably leads to misclassification, of short pauses as long pauses, and vice versa.

We have used the Maximum-Equalisation algorithm[15] to optimise separation of the pause duration distributions. Given an estimate of the threshold that best defines the boundary between the two distributions, we define the means and standard deviations for the pause duration distributions (in natural log), the mean and standard for the speech segment duration distribution defined by the long pause duration distribution (in natural log), and a number of other statistics.

Long pause duration

The fact that the long pause duration distribution is lognormal complicates interpretation. Because the distribution is lognormal, it follows that the underlying process reflects interactions involving a number of variables. A provisional list of candidates would include conceptualisation, planning, intention, attention, lexical search and syntactic formulation, although breathing too is likely to be a factor. Unlike modular models that rely on fractionation or decomposition, therefore, it is not appropriate to identify a single hypothetical process with the result. Thus, unlike modular approaches to language production, it is not our assumption that the procedure measures a single process. Rather, our approach assumes that the distribution statistics reflect interactions involving many or even all of the processes listed above.

14 *Ibid*.

15 McLachlan, G and Peel, D (2000). *Finite Mixture Models*, Wiley, New York.

Figure 4 depicts the impact of practice on mean Long Pause duration for each individual's first and second language. Each participant has contributed two points to the figure, one for her or his first language, English or Italian, and a second one for her or his second language, Italian or English. The figure shows that trend lines fitted to the two sets of data intersect at approximately log 3.9 (base 10) or 7,000 hours of practice.

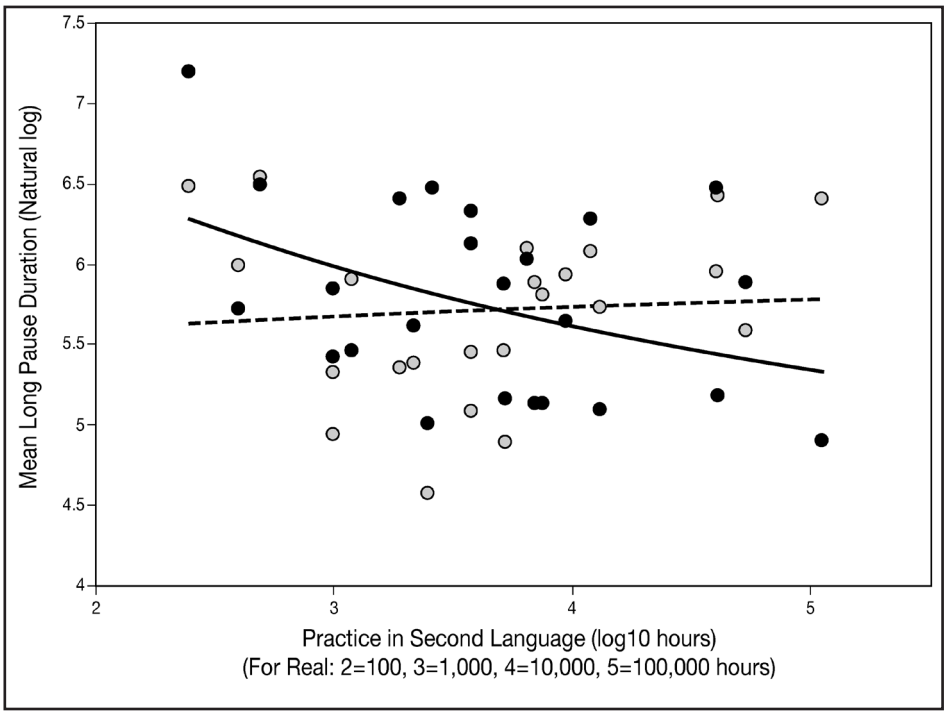

Figure 4 Relationship between practice and pause duration for L1 (open circles) and L2 (filled circles)

Discussion

Acquisition of second language skills

Our project is based on the premise that second language learning can be treated as a skill, and that measurement of practice is therefore critical[16].

Our first objective concerned the time required to attain fluency in English. Two parameters from a large potential suite suggest that people with Italian

16 Speelman, C and Kirsner, K (2005). *Beyond the learning curve: the construction of mind*, Oxford University Press.

First Language skills require at least 7000 hours to attain fluency in English. Assuming 50 hours of engagement in English per week (for an immigrant, not a second language student), this estimate rounds out to a minimum of three years.

Other studies provide even more depressing answers. Magiste[17], for example, used de-contextualised, as distinct from natural, language tasks and found that people who migrated from Germany to Sweden early in life did not reach the performance levels of native Swedes even after 20 years. But engagement is unspecified for Magiste's samples; some or all of them could have remained in German speaking communities.

Magiste's experiment included two groups of people, Swedish First Language Speakers (SFL, who had lived in Sweden from birth), and German First Language–Swedish Second Language Speakers (GFL-SSL, who had been born in Germany). She did not estimate number of hours of engagement in Swedish, and, as acknowledged by her, individual variation in this variable would have been considerable. Instead, she assigned each GFL-SSL participant to a group based on 'number of years in Sweden', and measured their performance on a series of de-contextualised tasks including Picture Naming Latency. The results were striking. Picture Naming Latency for the GFL-SSL group declined steadily as a function of practice, across the seven sub-groups that had been in Sweden for one to 20 years. The most significant result for our argument is that the performance levels for the GFL-SSL group never approached the values achieved by the Swedish First Language group, even after 20 years.

An even more depressing note was sounded by Thomas and Collier[18]. Children who entered the USA at age 12 showed steady gains in English throughout their secondary school years, however they actually fell further and further behind their English First Language peers throughout the same period. 'These findings show that there is no shortcut to the development of cognitive academic second language proficiency and to academic achievement in the second language. It is a process that takes a long, long time'[19].

Language differences and transfer between languages

Several authors have highlighted the role of similarity between a student's first and target languages. A significant part of the argument against the study of Mandarin, Cantonese and Japanese depends on the assumption that English First Language speakers find these languages significantly harder than languages that

17 Magiste, E (1979). The competing language systems of the multilingual: a developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behaviour*, 18(1): 78-89.

18 Thomas, W and Collier, V (1997). *School effectiveness for language minority students*. NCBE Resource Collection Series Number 9, Washington DC.

19 Thomas, W and Collier, V (1997). *Ibid*: 638.

belong to one or other of the Indo-European family of languages. The extent to which this analysis depends on the need to master reading and writing as well as speech is unclear. The value of transfer between languages depends on numerous factors including vocabulary, syntax, prosody and pronunciation. The assumption that languages such as Mandarin, Cantonese and Japanese are an order of magnitude harder to learn is insecure, and may depend on the detailed transfer patterns between the languages involved, and the level of mastery required.

Consider the following result from an unpublished study by a Japanese student visiting our laboratory. Australian students of Japanese were asked to read words from Hiragana (Japanese words) and Katakana (loan words). The results showed that Australian students experienced reading problems with Katakana, the script used to depict Japanese loan words. Ice cream or *aisukuri-mu* is an obvious example. The word is borrowed from English but, while it is easy to learn, it is actually difficult for English speakers to pronounce. Transfer patterns between English and other languages are not in all cases obvious. Our overall study will provide evidence on the comparative difficulty of reaching first language proficiency in a variety of languages.

It is important to note however that the US-based Defense Language Institute provides tuition for six to sixteen months, depending on language, and that the longer period includes languages such as Japanese, Chinese and Korean.

The relationship between fluency and integration

The questions designed to elicit information about integration were not included in the pilot study reported here. But the question goes to the heart of the objectives of the overall project; what is the role of language in integration and multiculturalism? Does integration simply follow the language function, for example as if the two processes are coupled? Perhaps language leads to integration so that integration is only possible after a certain level of English proficiency is acquired. When we have collected integration measures as well as fluency measures for people from numerous language and cultural groups, we will be in a position to address the broader issue.

Will new technologies or combinations of technologies transform the time taken to master second languages?

The most important issue raised by our study involves the sheer amount of time it takes to master a second language. Whether the precise figure is 7000 hours

or 70000 hours or it involves different amounts of time for different language skills is not the issue; it invariably involves a significant cost to the student. Yet the case for second language learning, and for a broad distribution of second language skills in the Australian community is overwhelming for educational, social, cultural and security reasons.

Perhaps surprisingly the security case does not involve terrorism, for recent events indicate that proficiency in a language provides no protection against extreme acts of hostility toward host communities. But these acts involve less than 1/10th of one per cent of the immigrant community, and cannot be used to shape *integration* policies.

A more appropriate challenge involves the facilitation of proficiency in English among the immigrant community generally, in a climate of limited resources, in regard to both teachers and funds. Four factors merit consideration, and, like most technical challenges, significant change is likely to involve interaction among several factors.

Virtual worlds

A recent article by Easteal[20] provides part of the answer. Technology has reached the point where it is possible to establish virtual environments between individuals and groups more or less anywhere in the world. Easteal, a legal expert from the University of Canberra undertook a 'virtual sabbatical' at Durham University in the UK, and participated in a wide range of social and academic events. I inhabit the same world when my London-based son calls me from his CISCO office. Virtual environments have barely been sampled for learning purposes. Given a camera on my head, a microphone on my lapel, and a transmitter on my back, I can include anybody anywhere in my everyday life.

Situational and contextualised learning

The second factor involves situated or contextualised learning. The trade world has honoured the apprenticeship model for centuries, with on-the-job training for electricians, mechanics and a host of other 'trades'. But over recent decades the medical profession has gradually embraced a variant of this model as well.

The Clinical Training and Evaluation Centre at UWA (ie, CTEC) provides an interesting example of the concept. Traditionally the medical model followed the trade model with supervised practice in an apprenticeship system. However, owing to decreasing patient availability, increasing demand for complex and hazardous procedures, and the rising number of trainees, the opportunity to become competent in a medical skill using traditional training methods

20 Easteal, P (2009). *Voices of the Survivors*, Spinifex Press, Melbourne.

has diminished. The solution for medicine involved the development of simulated environments including, for example, operating theatres. Simulated environments are designed to mimic the real thing while enabling minimal patient involvement during the early stages of skill acquisition, providing numerous opportunities for unlimited practice without risks to patients, and providing expert feedback to improve accuracy and reduce errors.

The language learning equivalent could place engineering students in China at the 'coalface' in Rio Tinto, Chinese Second Language students in Australian restaurants in Guangdong, and it could tailor the environment available to each student according to his or her preferences and needs, for cooking or system management in an iron ore mine for example. Importantly too, because such a system provides a form of social system, it has the potential drawing power associated with pen pals and even the Web dating world. While these opportunities involve risks, they also offer the type of motivation that fuels learning in all forms in the teenage community.

Collaborative learning

A third factor involves Collaborative Learning. The potential to create tailored groups is considerable. Groups would generally need a language expert to be 'on call' for planned support but many of the support skills could be met by individuals at more advanced levels of proficiency. Collaborative Learning, though, is not without its problems[21].

Conclusion

In conclusion, it is our contention that the critical issue involves transformation of the concept of a community. When I can share part of my daily life with a remote 'cousin' in Quangdong, and work and play in a remote community, I will in be in a position to master the language skills required to live and work in that environment. This is, of course, equivalent to living in a foreign country, but without the cost and time limits usually associated with international travel.

21 Roberts, T S and McInnerney, J M (2007). Seven problems of online group learning (and their solutions). *Educational Technology and Society*, 10 (4): 257-268.